AD 651 783

MOTION PICTURE TESTING AND RESEARCH

James J. Gibson

Army Air Forces
Washington, D.C.

1947

AD-651783

Motion Picture Testing and Research

REPORT NO. 7

Best
Available
Copy

Army Air Forces
Aviation Psychology Program
Research Reports

Motion Picture Testing and Research

REPORT NO. 7

Edited by
JAMES J. GIBSON
Associate Professor of Psychology
Smith College

1947

Preface

The research to be described in this report originated in the effort to utilize the motion picture medium for purposes of psychological testing and examining in the Army Air Forces. By extension, the original research came to be concerned with additional problems, such as the presentation of films to groups, the problem of effective instruction by means of projected pictures and films, and the problem of the representation of three-dimensional space by pictures. The problems encountered were relatively novel, and the results may be of interest to civilian psychologists, to educators, and to individuals concerned with films, as well as to aviation psychologists themselves. Training in the Army Air Forces made use of pictorial methods on a vast scale. The data to be reported on several aspects of this training are to some extent relevant not only to aviation but also to general education and to the theory of visual learning.

The studies made by the Psychological Test Film Unit were performed for concrete military purposes having to do with the selection and training of men for the Army Air Forces. The reports originally submitted were written with those purposes in view. The present account includes an attempt to reinterpret these data from a more general point of view and to point out their implications in other fields. Particularly in chapter 9 on the perception and judgment of aerial space, and in chapter 10 on the techniques of teaching with the film medium, a general and theoretical discussion has been included.

The accomplishments to be recorded were the product of the cooperation and the energy of a number of individuals. Although credit is difficult to assign, the names of those chiefly responsible are given at the beginning of each chapter. Special acknowledgment should be made of the contribution to all phases of the research of Robert M. Gagne, Assistant Director of the Unit.

JAMES J. GIBSON,
Lt. Col., Air Corps.

Santa Ana, Calif., February 1946.

Table of Contents

Chapter | Page

Chapter		Page
	The Problem of Testing—Types of Judgment Indicative of the Ability to Perceive Distance	196
	Judgments of the Relative Distance of Two Objects Side by Side	196
	Judgments of the Absolute Distance of a Single Object	196
	Judgments of the Size of a Far Object in Relation to That of a Near Object (Size-Constancy Method)	197
	Judgments of Absolute Distance During Locomotion	198
	A Photographic Test for Distance Perception	198
	The Relation Between Size Constancy and Distance Perception	198
	An Experiment to Validate the Photographic Method of Representing Distance	200
	The Experiment	201
	The Results	204
	Implications of These Results for the Theory of Aerial Space Perception and the Theory of Perceptual Constancy	209
	Construction of the Distance Estimation Test, Form A (CP212A)	212
	Construction of the Distance Estimation Test, Form B (CP212B)	217
	The Ability to Judge Distance and Space in Terms of the Retinal Motion Cue	219
	Types of Retinal Stimulation in Relation to Visual Motion Perception	219
	Retinal Motion Perspective	221
	The Effect of Eye Movements on Gradients of Velocity	224
	The Perception Resulting from Retinal Motion Gradients	226
	Motion Perspective When the Observer's Movement is not Parallel to the Terrain; Application to the Problem of Landing an Airplane	226
	A Motion Picture Test for Accuracy of Judgment During Landing	230
	Method of Constructing the Test	230

CHAPTER ONE

Historical Background of Motion Picture Testing and Research*

HISTORY AND ORGANIZATION OF THE PSYCHOLOGICAL TEST FILM UNIT

Establishment

The Psychological Test Film Unit was established in October 1943 at Santa Ana Army Air Base, Santa Ana, Calif., as a continuation, with specialized functions, of the Perceptual Research Unit of the Psychological Section, Office of the Surgeon, Headquarters, AAF Training Command, Fort Worth, Tex. Its primary purpose was to develop the work already begun on an experimental program of motion picture test construction and on allied problems involved in the psychological use of films. It was located near Hollywood in order to further this purpose. The unit was attached to the Station Hospital, Santa Ana Army Air Base, and in order to simplify military administration the enlisted men of the unit were drawn from Psychological Research Unit No. 3 and, in this period, were included in the military organization of that unit.

The technical direction of the unit was, like all the original units of the AAF Aviation Psychology Program, the responsibility of the Surgeon, AAF Training Command. On 1 November 1944, following the transfer of Santa Ana Army Air Base to the Personnel Distribution Command, the officer personnel of the unit were transferred to Headquarters, AAF Western Flying Training Command. Ten enlisted men on duty with the unit were also transferred to that command and were placed on detached service at Santa Ana Army Air Base. Special arrangements were made to permit the unit to continue to be housed at Santa Ana Army Air Base. In October 1945, the officer and enlisted personnel of the unit were transferred to Headquarters, AAF Central Flying Training Command, Randolph Field, Tex., but remained on detached service at Santa Ana Army Air Base.

Prior to the establishment of the unit in October 1943, the production of films for motion pictures had been carried out by

*This chapter was written by the editor.

arrangement with the Motion Picture Branch of the Technical Data Laboratory at Wright Field, Ohio. After that time films for all aptitude and achievement tests were produced by the AAF Motion Picture Unit, Culver City, California, the producing studio for training films in the Army Air Forces. The services of this agency for the specialized purpose of psychological testing were authorized by the Office of Motion Picture Services in Headquarters, Army Air Forces in Washington, and by arrangement with the Training Aids Division (Office of Assistant Chief of Air Staff for Training) in New York, which had charge of training films. Production of the aptitude tests was carried as "Special Film Project No. 112." The achievement tests in aircraft recognition, to be described, were classified as training films and each test was given a training film production number. Other proficiency tests were produced individually as special film projects.

Nearly all film production, and consequently the trying out of experimental tests, was carried out on standard 35-mm. film, sound being added to the film at a later stage just before completion of the test. The standards of production were those of the large commercial studios. Film editing and cutting was performed partly at Culver City and partly at Santa Ana. Preliminary experimental photography and experimental administration of tests on 16-mm. film without accompanying sound would have been desirable. Such photography does not require highly-trained professional technicians and could have been carried out by the Psychological Test Film Unit itself. This was, however, impossible during 1944 and most of 1945 because of the shortage of the necessary equipment. After obtaining 16-mm. camera equipment in September 1945 some experimental photography was performed at Santa Ana.

The administration of experimental tests to groups of preflight students was coordinated with that of Psychological Research Unit No. 3 when that unit was at Santa Ana Army Air Base. Close cooperation existed between the two units with respect to test construction and exchange of services. When the Psychological Test Film Unit was first activated, the continuation of the research on printed perceptual tests originated by the Perceptual Research Unit at Forth Worth was transferred to Psychological Research Unit No. 3 and was carried out by that unit after October 1943.

Administration of experimental forms of aptitude and proficiency tests to local groups was no longer possible following the removal of the Preflight School from Santa Ana Army Air Base in November 1944. Consequently the major portion of the work of the unit following that date was devoted to the design and pro-

duction of motion picture tests. Preliminary experimentation to determine type and difficulty of items was carried out on small groups of enlisted men at Santa Ana Army Air Base, but administration for statistical analysis had to be conducted at psychological units located at centers of training activity.

Personnel

All members of the Psychological Test Film Unit were engaged primarily in research, although various members specialized to some degree in film editing, still and motion-picture photography, art work and drafting, film projection, statistical analysis, test construction, training experiments, the writing of scripts, and the composing of research reports. The following individuals were the relatively permanent members of the organization:

Rank	*Name*	*Army Serial No.*
Lt. Col.	Gibson, James J.	0-901965
Capt.	Eisenberg, Ralph M.	0-562936
1st Lt.	Gagne, Robert M.	0-563064
T/Sgt.	Bornemeier, Russell W.	13076283
S/Sgt.	Finney, Ben C.	39842343
S/Sgt.	Wiegand, Edward C.	16077369
Sgt.	Classer, Nathan M.	32931276
Sgt.	Lamkin, Hibbard	35529040
Sgt.	Schafer, Alfred H.	39034559
Cpl.	Bannerman, Edward M.	39198852
Cpl.	Borin, Leighton H.	17098303
Cpl.	Slater, Gerald M.	39552035

The following individuals were members of the organization for shorter periods of time:

Rank	*Name*	*Army Serial No.*
Major	Lehner, George F. J.	0-1000257
2d Lt.	Orvis, Clay H.	17030003 (as E/M)
2d Lt.	Tice, Frederick G.	33528167 (as E/M)
S/Sgt.	Price, Hubert B.	19200376
Pfc.	Hackbarth, Edward A.	36746950
Pfc.	Luft, Joseph	33537710
Pfc.	Marion, Arthur J.	16111664

Mission

The mission of the unit may be summarized as follows:

(1) The construction of motion picture tests for classification and other purposes.

(2) Research on motion picture and photographic methods of psychological testing and on problems of administering and scoring motion picture tests.

(3) Research on training problems connected with or amenable to motion picture and photographic techniques as directed.

(4) Advising, editing, and cooperating in the production of films for test development in any units of the AAF Aviation Psychology Program.

AREAS OF RESEARCH

Research Objectives

The aim of research, generally, was the establishing of facts which might lead to improved methods for classification and training of AAF personnel. The area of facts studied was one which, although it cut across many fields of psychology, had one underlying problem common to all of them—the problem of what might be called the perception of pictures. Pictures, of which motion pictures are the most realistic form, offer unexplored opportunities for the development of psychological tests and for devising methods of training. Pictorial or photographic tests had already demonstrated their usefulness, and presumably motion picture tests might have even greater potential value; pictorial and motion picture methods of training were likewise known to be highly effective and to be in need of study. The AAF had adopted for its training program a policy of "visual education" and it was therefore important to look for the psychological principles underlying this rather vague slogan.

Assumptions and Hypotheses

Apart from the specific assumptions involved in particular projects, the guiding hypotheses of a research program are frequently not explicitly formulated. It is useful, however, to attempt to do so. The following represent a number of general hypotheses governing the research of the Psychological Test Film Unit:

(1) Tasks performed by military fliers and other members of the aircrew are to a large extent visual in nature, and, consequently, tests of visual and spatial perception are important for measuring aptitude for aircrew duties.

(2) The perceptual discriminations required in flying can usefully be studied by a combination of psychological methods: first the making of job analyses of the task of the flier, and second breaking down these descriptions in terms of the abstract theory of perception.

(3) The methods of the psychophysical experiments are valuable for the study of complex visual discriminations, and are in some ways better techniques for the construction of motion picture aptitude tests than are the conventional methods used in constructing printed tests.

(4) The devising of motion picture tests should be confined to those functions for which the medium is uniquely adapted. (The unique characteristics of the motion pic-

ture medium for psychological testing are discussed in the following chapter).

(5) The kind of behavior primarily involved in the task of flying is *locomotion in space* and on that account is extended in time. Hence, the performances required in flying are predominantly characterized by motion, by being continuous, and by possessing tempo. Therefore, so far as group testing is concerned, the motion picture medium should have a number of unique advantages for measuring the non-academic aptitudes and abilities peculiar to flying over the material reproducible in printed tests.

Areas of Research

The Psychological Test Film Unit was a research organization which constructed tests but did not administer and score them for purposes of selection and classification of students. The areas of research within which tests could be constructed were theoretically unlimited. Any test which could be presented on film, whether categorized as perceptual, intellectual, or emotional, could have been attempted if the test were needed. Requests for tests, or the filming of partially-completed tests, were accepted from other units of the AAF Aviation Psychology Program having specialized fields of research. Tests were devised or partially developed in such fields as comprehension, resistance to confusion, and performance under emotional stress, but the difficulty of such research, the problem of adaptation to a medium still being explored, and the limited experience of personnel in such fields of research, combined to make progress in these areas slower than in the case of perceptual tests.

The majority of the tests produced were therefore perceptual in nature since motion pictures have their most obvious application in this field, and since the work of the unit was originally concerned exclusively with such tests. It was also concerned with a technique or method of testing, the motion picture, possessing its own characteristics and its own special opportunities for psychological investigation. In general, the research of the organization may be divided into general parts: Test Development, Training Research, and Problems of Technique.

Test Development

The most important research objective of the Film Unit was the construction of motion picture tests for aircrew classification purposes. The general procedure was to formulate a hypothesis regarding a function thought to be valid for prediction of success in training in one or more of the aircrew specialties, i. e., pilot, navigator, or bombardier. The test was next put together, accord-

703317—47—2

ing to methods described more fully in a following chapter, in the effort to measure the desired ability. The experimental test was administered to a large group of aviation students either while they were taking the official classification tests or during an early phase of their training. From test scores thus obtained the reliability of the test and the intercorrelations with the tests being used for classification purposes were ascertained. The validity of the test was determined by correlating the test scores with success or failure in later phases of aircrew training. If the intercorrelations of the test with the classification tests were low enough and the validity of the test high enough to add significantly to the prediction of the classification battery, the test was considered for inclusion in the classification battery.

Detailed descriptions of individual tests are to be found in a later chapter. However, the general areas in which tests were constructed are as follows:

1. Tests of ability to judge motion and locomotion. Three tests were constructed requiring judgments of visual motion and one which required judgments of one's own motion during simulated flight. The last of these involved the isolation, control, and representation of the visual cues believed to be primarily the basis for landing an airplane. This test simulates the view of the field during a landing glide. The film was produced by miniature photography.

2. Tests of ability to judge distance. One test was completed in this area, after considerable research and the constructing and trying out of a number of preliminary tests.

3. Tests for orientation in space. As a result of trying out several pictorial types of test, and preliminary research with motion picture forms, one test was completed for the ability to maintain orientation in the air after a series of turns, and another for orientation in the traffic pattern.

4. Tests of ability to perceive slight movement. Since the ability to synchronize a bombsight depends on this function, two tests were constructed in this area after an attempt to isolate the relevant perceptual variables discoverable in the job.

5. Tests requiring multiple perception. Two tests were produced with the aim of measuring a phase of "alertness" which was repeatedly emphasized in job analyses of the pilot's task. This phase appeared to be the ability to take note of multiple simultaneous stimuli (such as instruments) or to keep in mind a number of diverse sets to respond to these stimuli. Psychological theory in the field of attention being inadequate, however, the analysis could not be clarified. The tests were finally constructed as copies of flying situations, only partly abstracted from reality. They involved reactions to a schematic instrument panel and the

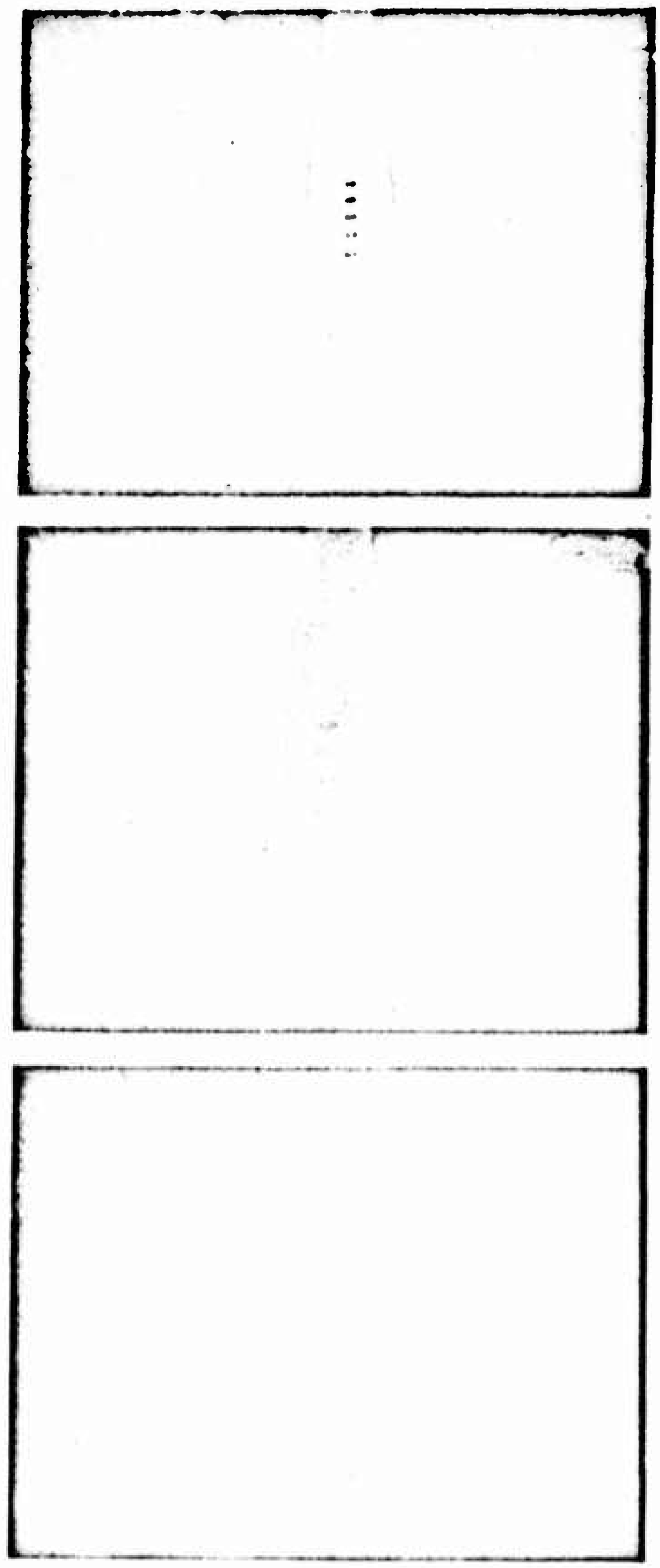


FIGURE 5.3.—Successive Appearances of the Runway During a Landing Glide.

functions were loosely called "flexibility of attention" and "integration of attention."

6. Tests involving sequential perception. The psychological aptitudes or functions involved in the perception of continuous changing situations are likewise not well understood. Two tests were constructed, however, on the basis of an abstract stimulus situation. They involved the ability to put together successive partial impressions into a complete visual figure. In one test the figure to be perceived was traced by a moving spot; in the other it was exposed, one part at a time, through a slot which moved across it.

7. Tests of perceptual speed. One test was constructed in this field. It required the subject to perceive and reproduce a visual pattern exposed for a very brief interval. The motion picture method makes possible the control of exposure intervals in group testing.

8. Tests of comprehension. An aptitude test was devised which was analogous to reading-comprehension tests except that the material was presented by film. The ability involved, therefore, was comprehension of audio-visual material of the sort presented in demonstrations and lectures rather than of the written material of textbook instruction.

Training Research

The mission of the Psychological Test Film Unit included, in addition to construction of motion picture tests for classification purposes, "research on training problems connected with or amenable to motion picture and photographic techniques." The two major projects in training research were, first, the construction of proficiency or achievement tests in motion picture form and, second, the evaluation of training procedures.

Proficiency Tests. Two proficiency tests, in navigation and bombardiering, were undertaken by the Film Unit in collaboration with research units of the AAF Aviation Psychology Program and were constructed to meet their requirements. These were the Navigation Proficiency Test (Map Reading and Dead Reckoning) and the Target Identification Test. In addition, the training research on aircraft recognition conducted by the unit demonstrated the need for two motion picture proficiency tests for evaluating such training. These tests are the Aircraft Recognition Proficiency Examination, Forms A and B, and the Aircraft Recognition Proficiency Examination for Flexible Gunners.

Evaluation of Training Procedures. The two major projects in research on the evaluation of training procedures were, first, a comprehensive study of aircraft recognition training, and second, an evaluation of the effectiveness of Army Air Forces training films.

In both undertakings the essential problem was that of determining what methods or techniques of training (of the alternatives which were practicable) were the most effective in producing learning by the trainees. Both studies necessitated the use of a criterion for learning, and therefore involved the construction of achievement or proficiency tests. The experiments carried out consisted of the setting up of a control group of aviation cadets who were trained in the ordinary way (or, on occasion, were given *no* training with the technique being investigated), and one or more experimental groups who were trained by a different method or methods. The experimental and the control groups were either drawn from the same population in large numbers, or were matched for proficiency by a pretest. At the end of the training period, both groups were given the criterion test and any differences in proficiency were evaluated.

Problems of Technique

In addition to the above fields of research, a more or less continuous study was made of the unique characteristics of the motion picture as a device for psychological testing and for training. A number of novel problems arose in connection with the representation of visual stimulus situations by the method of taking and projecting motion pictures. A number of other problems were studied having to do with the group test situation in which a common stimulus object, the motion picture screen, is present to all testees in the room but is seen at different distances and angles of view by different individuals in the group tested. The results of these studies are reported in the appropriate sections later in this report.

CHAPTER TWO

The Use of Motion Pictures in the Design of Psychological Tests[c]

A motion picture film including sound can be produced only by the application of professional technical skills and at some expense. With respect to difficulty and expense of production, motion picture tests stand between printed tests and individual apparatus tests. With respect to the problems involved in test administration, motion picture testing is simpler than individual testing, since it can be carried out with groups, but in comparison with ordinary group testing, it involves special problems such as the operation of a projector and partial illumination of the testing room. For these reasons, it is important to design motion picture tests *only for such functions or abilities as cannot be adequately measured by printed tests.*

In order to utilize film effectively in psychological testing, the characteristics of the motion picture medium must be understood. This medium makes possible a novel way of presenting problems, events, situations, and stimuli. It enables the psychologist to put to his group of observers tasks which would be impossible to present by printed or pictorial methods, and even ones which could not be presented by an apparatus, or a miniature situation, without prohibitive expense. This chapter will attempt to describe and exemplify the unique characteristics of the motion picture medium for psychological purposes.

For use in the Aviation Psychology Program, final forms of 15 motion picture tests were produced, of which 12 are aptitude tests and 3 are proficiency tests. Each of these films had at least one and in some cases as many as four experimental forms from which selection of items for the final form was made. The experience gained in the design and construction of these tests by personnel of the Psychological Test Film Unit, while incomplete in many respects, is sufficient to permit a formulation and summary of the properties of motion picture testing which distinguish it from other forms of testing.

[c] This chapter was prepared by Robert M. Gagne, Russell W. Bornemeier and the editor, with the assistance of Leighton H. Borin.

INTRINSIC AND UNIQUE CHARACTERISTICS OF THE MOTION PICTURE MEDIUM

For the purpose of test design, motion pictures may be considered to offer four major possibilities which are unique in contrast to printed tests. These are movement, sequence, pacing, and realism. Any of these characteristics may be shown to be an essential factor in the construction of a particular test. Although they are admittedly interrelated, they will be described in order.

Movement

The obvious characteristic which may be possessed by the stimulus material of a motion picture test is movement. It is important to note that two different and distinct kinds of movement-perception may be induced on the screen, movement of *objects* and movement of the *observer himself*. Particularly in the latter case, the motion picture yields an enhanced perception of the three-dimensional quality of the space portrayed on the screen.

The representation of one or more moving *objects* can be employed whenever the function to be measured requires visual judgments of the speed or direction of motion. In Speed Estimation Tests CP205B–I, CP205B–II, and CP205B–III, for example, the aspect of movement which is to be judged is *velocity*. In Speed Estimation Test II, five different velocities are represented which must be remembered and distinguished from one another. In Speed Estimation Tests I and III the judgment of velocity is obtained indirectly by requiring an estimation of the point of coincidence of a moving object with (a) a stationary event, and (b) another moving object, respectively. There are, of course, many other possibilities of obtaining judgments of velocity which have not yet been employed in motion picture tests, some of which are indicated in the latter part of this chapter. In the case of the Minimal Movement Test CP213C and the Drift Direction Test CP221B, the effort is made to measure thresholds for the perception of motion indirectly, by requiring judgments of the direction of movement. All of the tests mentioned above, and other tests referred to in the discussion which follows, are described in chapters 5 and 6.

A second and equally important type of movement which may be represented in motion pictures is *motion of the observer*, i. e., of the examinee. By putting the camera in place of the observer and moving it through an artificial or a natural environment, the experience of visual motion is imparted to the observer himself. Since a large part of the task of learning to fly may be said to consist in learning to judge and control locomotion in a new kind of space, or at least a spatial world which is new to the beginner, the useful-

ness of this type of test for the selection of pilots should in theory be great. An example of a test which utilizes this characteristic of motion pictures is the Landing Judgment Test CP505E, which was designed to measure the individual's ability to learn to make a certain type of judgment required in landing a plane, namely a judgment of the point at which the plane's glide is aimed.

Other examples of the utilization of the movement of the observer may be found in two additional tests by the Psychological Test Film Unit, the Flying Orientation Test CP107A and the planned but unfinished Following Maneuvers Test CP532A. The Flying Orientation Test represents the travel of a plane over the ground in an artificially complicated flight path, as might be seen by an observer looking down from the plane; the Following Maneuvers Test provides the visual experience of being in a plane looking forward and viewing the ground and horizon during certain simple maneuvers. Although the latter test was not completed, a comparison of its characteristics with those of a similar printed test is significant for the present discussion.

The Following Maneuvers Test is intended to measure the ability to acquire the "schema" of the control movements of an airplane in response to cues from an unfamiliar spatial framework. These responses are necessary for controlling the flight of an airplane. The movements of horizon and terrain on the screen represent the visual experience which the examinee would have if he were in a real plane performing simple maneuvers. In response to the view presented, the subject records the appropriate movements of stick and rudder on a special overprinted answer sheet on which these responses have been schematically represented. A printed test, proposed by Psychological Research Unit No. 3, a revised form of Flight Orientation Test CP528A, is in many respects similar to the motion picture test, particularly since both utilize the same responses and method of recording. In the printed form, each item consists of a sequence of stationary views, printed from left to right, of the horizon and terrain, as seen from the cockpit in successive stages of movement through the maneuver. It is apparent that the situation presented by the printed form is one that requires the examinee to *interpret these still pictures as moving*, whereas the motion picture test enables him to perceive the motion directly.

Since the types of movement presented in motion picture test items can include synthetic motions by the use of special effects photography and frame-by-frame photography, the test designer has at his command a rather complete *control* of the movement stimulus. He can plan and translate onto film any kind of motion he wishes to present, including both natural and artificial locomotion of the observer.

Sequence

A second type of unique material which may be presented by motion pictures consists of items within which a temporal sequence of events must be perceived by the examinee. The sequence may be discontinuous, in which case a series of discrete stimuli are shown in orderly succession, or it may be continuous, in which case the parts of the sequence change from the beginning to the end of the item presented. In the latter case, although movement is present, it is not the factor which is being judged.

An example of the presentation of a *discrete sequence* is a test of immediate memory span for digits or verbal material, in which elements are presented in successive controlled exposures and in which the order of the elements is important. Measurement of this sort is ordinarily done in the psychological laboratory with the use of various exposure devices. The motion picture film and the projector simply afford the possibility of standardized presentation of such tests to groups of individuals.

Items which have the characteristic of *continuous sequence* are particularly adaptable for motion picture presentation. Two examples of this kind of item are to be found in the Successive Perception Tests CP509C–I and CP509C–II. In the former, a black and white pattern is seen, a small part at a time, through a slot which moves continuously from the top to the bottom of the screen. The task is one of perceiving the total pattern, whose parts have been seen in successive impressions, and identifying it from five alternative patterns simultaneously presented. In Successive Perception Test II, a spot moves through an irregular path on the screen, and the pattern made by the total path must be visualized by the examinee and chosen out of five alternatives. Many variations of the relatively simple items of the tests just described may, of course, be obtained. As in the case of motion, the characteristic of sequence may be necessary for a test in which the attempt is made to reproduce the essential stimulus conditions of the job itself. Many tasks performed by aircrew members include the element of sequence. Of particular importance is the ability to learn a *procedure,* in which a number of acts must be carried out in a particular order. The general function involved in such tests might be said to be the comprehending of temporal rather than spatial patterns.

Pacing

A third unique characteristic of motion pictures is pacing, or the tempo of the elements presented. Broadly defined, pacing implies control over all temporal aspects of the test material. In the case of any particular test, temporal control may be applied to the stimulus-interval, to the response interval, or to both. In some

cases it may be desired to have the test duplicate as exactly as possible the pace or tempo of the job itself; in others it may be necessary to abstract from the job situation some psychological function best measured by items which do not follow the temporal characteristics of the job itself. An artificial pace can be set by the test designer.

Pacing Which Duplicates That of the Job. For the construction of proficiency tests, the element of pacing may make the difference between an adequate and an inadequate measure. For example, the job of pilotage requires the navigator to locate his position on a map by observing terrain which is continually moving, and on which any particular reference point can be seen for only a few minutes. In this situation the navigator is reacting to a sequence of events in time, a sequence which would be destroyed by breaking up the material in order to form items. Presenting this sort of test by means of still photographs, even in a time-limit test, fails to duplicate the paced nature of the navigator's task. The only accurate means of representation of paced tasks of this nature is the motion picture, or a device which possesses its characteristics as do some "trainers."

An example of a test which presents the same pacing as that of the job itself is the Navigation Proficiency Test (Map Reading and Dead Reckoning) in which the tempo and the sequence of events were made to correspond closely to those of an actual navigation practice mission. It is, of course, impossible to have "items" of the usual sort in a test such as this without interrupting the sequence and disturbing the natural pace of the navigator's task. Consequently, in the case of the test described, measurement is accomplished by requiring the examinee to keep a log just as he would in flight, which is later objectively scored by hand.

Artificial Pacing. Various degrees of departure from the exact duplication of the temporal aspects of the job may be employed in devising a test. One possibility is exemplified by the Flexibility of Attention (CP411E) and Integration of Attention (CP415A) tests, in which the pace is speeded by requiring that a number of disparate events, occurring at irregular intervals, be observed and recorded by the subject within 12-second phases. The natural irregularity of the temporal sequence of such jobs as keeping watch over a complex instrument panel is maintained in these tests, but the increased pace of the test makes possible a large number of responses within a limited period of time, and simulates the pace of the job itself *in times of emergency.*

A second example of a test in which the natural pacing of the job is altered to permit greater scoring possibilities is the Target Identification Test. In this case the sequence of events in the approach of a bombardier to the target is interrupted by a num-

ber of stop-frame shots of the terrain ahead, each with a superimposed lettered grid which makes it possible for the examinee to note the location of the target *if he has identified it*, and then to record the appropriate letter. While the pacing of the approach to the target is duplicated, the introduction of stop-frames of the terrain makes possible the measurement of various degrees of success in identifying the target for which the bombardier has been briefed.

In the Landing Orientation Test CP106A the examinee is required to learn and remember the appearance of a landing field at a specific point in the traffic pattern. The specific point to be learned and remembered is signaled by voice recorded on the sound track. After having observed this point during a normally-paced flight in the traffic circuit, the subject is required to identify it as falling within one of five successive intervals, A, B, C, D, or E, which are announced during a repetition of the same flight. It will be noted that an essential characteristic of this technique is the presentation of the alternative responses sequentially, rather than simultaneously, which permits the natural pacing inherent in the job itself to be maintained in the test situation. The discrimination required in this test is one involving the *right moment at which something is to be done*. In flying, as in any performance which is extended in time, the choice of the right moment for action in adjustment to a changing situation is a vitally important part of the job, and aptitude for this type of judgment can scarcely be measured except by reproducing the changing situation with film. The ability to react at the right time is, in fact, very close to what is often meant by the general term "judgment" in flying.

Controlled Pacing of Items and Responses. Some types of tests require arbitrarily controlled presentation-time of both items and response periods. Such control is possible in motion picture tests. However, the greatest importance of controlled speed of item presentation lies in its use for measuring the rapidity of any one of a number of perceptual and intellectual functions. The requirement of a definite speed of response for each item individually is of unique usefulness in designing tests which measure speed of reaction, or performance under pressure of speed.

1. Control of Interval for Item Presentation. An example of the manner in which speed of item presentation may be controlled is provided by the Plane Formation Test CP805C which was designed to measure a kind of perceptual speed. Each item of this test consists of a brief exposure of a pattern made up of five small airplanes. The subject is required to reproduce this pattern by marks on his answer sheet during the response period. A printed perceptual test can be "speeded" only by imposing a time limit on the items as a group and measuring speed by scoring the number

of items completed. The above test imposes a time limit on each item individually. All subjects have an equal chance at every item. Since the exposure interval is shortened progressively throughout the test, quickness of perception is measured directly.

A second example where a minimum speed requirement on each item is an important factor is the Aircraft Recognition Proficiency Examination. It was believed by the authorities responsible for aircraft recognition training and by the Film Unit that the job of recognition could be most realistically represented by short motion picture views of airplanes *in flight*. This method was in contrast to the existing practice of presenting either motionless views for split-second intervals or else silhouette views for indefinitely long intervals. In a sense, therefore, this test attempted to duplicate the pacing of the job itself. Views of the aircraft to be recognized were presented, in motion, for intervals in the neighborhood of 2–4 seconds.

2. Control of Interval for Recording Responses. In the majority of motion picture tests to be described in this report the interval for recording responses on standard answer sheets lasts for 5 to 8 seconds, or long enough to permit even the slowest subject to mark his answer. The determination of the length of this interval was obtained by experiment. The possibilities for testing which would result from shortening this response time have not been fully explored.

One possible use for speeded response intervals would be in a test which measures ability to perform some intellectual or perceptual task under the pressure of limited time. The ordinary speeded printed test undoubtedly induces this feeling of "pressure" in the subject; however, the motion picture test makes possible the inducing of this "pressure" on every single item. It is possible, therefore, that a more consistent measure of ability to perform a particular function under pressure may be obtained by a motion picture test. However, in order for such consistency to be achieved, it is necessary to set the length of the response period carefully. If the period is too short in relation to the difficulty level of the items, the "pressure" may become frustration, and other factors such as emotional confusion may be added to the situation to complicate the measurement of the desired function.

This suggests, however, a second and perhaps more useful possibility of measurement which results from the shortening of the time allowed for responses, i. e., the function which may be termed "resistance to frustration" or "efficiency under stress." The stress would be provided by response periods so short that they make it impossible for responses to be made and recorded. In the case of such a test, too, the regulation of the response periods in relation to the difficulty of the items must be carefully undertaken, for if

they are too short they will cause the subject to give up the task entirely. This result is not desirable, since it does not permit discriminative measurement of the desired function among the subjects (although it may be desirable to discriminate between those who do and those who do not "give up"). So far as possible, examinees should be maneuvered into an attitude of perseverance at the task throughout the entire test, so that any decrease in efficiency produced by discouragement or frustration will be differentially revealed in the scores. The individual tasks, then, must not appear insoluble, but at the same time they must actually be so difficult that genuine discouragement or frustration is produced. It is necessary, moreover, to demonstrate that a low degree of correlation exists between the speeded and unspeeded task before the hypothesis can be accepted that speeding produces an independent stress effect.

Self-Paced Tests and Unpaced Tests. A self-paced test is one in which the correct response to one item automatically produces the presentation of the next item. The ordinary printed test is not self-paced in the precise sense of this definition; it is not paced at all. Although the individual may adopt the fastest tempo of which he is capable in attacking each item, he can voluntarily omit items, and his proceeding from one item to the next is not determined by the correctness of his response. A truly self-paced test must be an individual test, e. g., an apparatus test, an example of which is the SAM Complex Coordination Test CM701A. A motion picture test for testing groups of subjects cannot be self-paced.

The natural result of the construction of the usual motion picture, whether or not a psychological test, is that it possesses a *specific* tempo or pace. In fact, an important contribution which motion pictures can make to a testing program, as previously indicated, is that of making possible the presentation of externally-imposed pacing, whether this be the "natural" pacing of the task, or some "controlled" pacing imposed by the test constructor to measure some particular function.

Realism

Although the objects and events presented in motion picture form are not fully "real" they approximate reality more nearly than do ordinary photographs and pictures or verbal descriptions. By moving, they become *animated* i. e., alive. This tendency for scenes on the screen to appear real should not be overlooked in designing certain types of tests. To a great extent, the observer loses himself in the scene, i. e., locates himself in the environment and in the situation being portrayed. This attitude of "being there and seeing it happen" is compelling; it can only be overcome by

deliberately attending to the frame of the screen image or to objects in the projection room. This tendency to adopt the attitude of reality is much more striking in motion pictures than in any other form of pictorial or photographic representation.

Participation by the observer in the situation being portrayed can be enhanced by a number of camera techniques. The location of the camera in the scene photographed can be such as to make the observer see what a participant sees. For example, one has a greater tendency to experience what the paratrooper experiences as he makes his jump if the camera looks out and down through the door of the plane than is the case if the camera takes the point of view of a mere onlooker or observer. If the camera moves and shifts its view appropriately, the onlooker can be made to identify himself with an active participant in the situation and to do what he does, firing at enemy fighter planes or even fighting in hand-to-hand combat. This use of the camera as a participant is in contrast with its more frequent use as a story-telling agent in entertainment films.

This characteristic of the motion picture medium is probably of most importance for test construction in the field of personality and emotion. Provided a high degree of participation by the examinee in the scene can be assured, his reactions to a variety of potentially emotional situations can be measured both directly and indirectly. As with other types of tests in this field, the outstanding difficulty comes from the necessity of insuring participation in the stimulating situation. In tests for measuring the reaction to induced emotional states, for example, the usual procedure is first to apply the emotional stimulus and immediately thereafter to measure the amount of change in some unrelated intellectual or motor activity. In some cases, as in the Steadiness Under Pressure Test CE211A, the emotion-inducing stimulus is given simultaneously with the measurements. However, in both instances, there exists the possibility that the subject will be able to dissociate the two activities or to concentrate on the task and disregard the distraction, and this fact is probably the greatest obstacle to measurement of the expected effect.

One method of overcoming this difficulty, at least in large measure, is made possible by the use of motion pictures. The emotion-inducing stimulus may be a complex scene in which at the same time a high degree of participation is induced in the observer. Since the presentation is relatively complex, the efficiency of a variety of types of performance, such as observation, distribution of attention, or memory, may be tested by utilizing the content of the scene itself. This method, therefore, accomplishes two things: (a) it facilitates the participation of the examinee in the scene, since it is impossible to perform the task (of observation, for

example) without watching; and (b) it prevents the examinee from dissociating the emotional stimulus from the task, since both are given by the same material. The method outlined has not been employed in test construction, nor have all its possibilities been experimentally explored. The usefulness of this characteristic of motion pictures appears to be promising in the field of personality test construction, the tendency to participate in the motion picture scene is a fact which could usefully be exploited in work with projective methods and in efforts to measure social attitudes and social perception.

CHARACTERISTICS OF MOTION PICTURES WHICH PERMIT CONTROL OF THE GROUP TEST SITUATION

Control of the Instructions for the Test

Instructions may be presented in three different ways with motion picture tests: they may occur simply as printed instructions appearing on the screen, as printed instructions with accompanying voice (sound track), or they may be presented only by voice and supplemented by explanatory action on the screen.

The second method of presentation resembles that of printed tests, when the administrator reads the instructions along with the examinee. For many types of test, the third type of presentation, voice accompanied by action on the screen, can be extremely useful. This is particularly true when the function to be measured involves motion, or complex types of visualization, or changing events. Such tasks can be presented more rapidly and understood more easily when animated or action scenes are shown accompanied by voice than when printed instructions are employed. Some of the tasks presented in motion picture tests would be difficult to explain in any other way.

The instructions presented orally through the sound track are, of course, perfectly standardized with respect to wording and emphasis for each administration and for every copy of a motion picture test. The diction and enunciation of instructions given by sound track are usually of high quality, since they are spoken by carefully rehearsed professional narrators. In the case of motion picture tests, furthermore, it is possible to repeat, change, or emphasize certain portions of the instructions at appropriate intervals during the test itself. This can be done without seriously interrupting the continuity of the test, since all subjects are at the same stage of progress. Such a device is useful for insuring understanding of the task when a modification is introduced.

Test Administration

Since complete instructions for taking a motion picture test may be included in the film itself, the duties of the test administrator

are materially reduced. Time limits are automatically standardized without use of a stop watch and without risk of error on the part of the test administrator. Because of this standardization, some of the less well-defined effects of the examiner-examinee relationship are also excluded. It is impossible to say in general whether this factor is or is not advantageous in administration, since it may vary with the particular test employed.

Since the items of a motion picture test are projected on the screen, and since the tempo of the test is the same for all, the tendency of examinees to try to observe the progress of his neighbors is absent, and the risk of cheating is reduced. There is still, of course, the *temptation* to look at another's answer to a specific item when one is unsure of the correct response, but the interval for recording responses (in the tests already constructed) is such that it is nearly impossible to do this successfully. After any brief period of uncertainty, there is not enough time to locate and copy the response from a neighbor's answer sheet before the next item appears. Experience in administering motion picture tests leads to the conclusion that *if they are properly paced,* the problem of copying answers is minimized or eliminated, even when subjects are seated side by side in a testing room without individual cubicles.

Item Presentation

The pictured item of a motion picture test always occupies the same position. It cannot be viewed sideways or upside down; in fact, it is removed from the control of the examinee. In some tests, such as one which requires orientation of a map, this may be undesirable. In others, chiefly perceptual tests, control of image position may be essential to accurate measurement.

Since motion picture tests present items in a temporal sequence, the subjects are prevented from selecting their own order of items, and from turning back to complete unfinished work. This type of control may be desirable or undesirable in designing a test.

FUNCTIONS PARTICULARLY AMENABLE TO MOTION PICTURE TESTING

In this final section, an attempt will be made to point out some of the capacities, aptitudes and proficiencies which can be measured by motion picture tests but which cannot be measured as readily by other means. A number of potential tests will be described, as examples of these functions.

Many of the characteristics of motion pictures, especially sequence and pacing, can obviously be possessed by individual apparatus tests. An apparatus which is supplemented by motion picture projection, or its equivalent, could exploit the characteristics of motion and realism. Some of the more elaborate synthetic

trainers in use actually do so. The advantage of motion pictures for the test constructor, therefore, is chiefly that of group testing as against individual testing. In the discussion which follows, it should be understood that the principal contrast which is being made is between motion picture tests and printed tests.

The functions which will be exemplified are directly or indirectly related to the characteristics of motion pictures, as distinct from other modes of presentation, which have already been discussed, namely motion, sequence, pacing, and realism. These characteristics presumably have their psychological counterparts. Human behavior, and the capacities latent in it, also involves motion, order, tempo, and the experience of reality. It is reasonable to suppose, therefore, that the motion picture makes available to the test designer not only a special method of measuring known factors of human ability but also gives him access to new and unnamed functions not accessible to conventional methods of test construction.

The functions which follow are listed in categories that are not presumed to be independent. They are listed as they are for reasons of convenience of discussion. The functions are as follows: discrimination of visual motion and locomotion; perception of space and distance, particularly during flight; maintaining orientation during locomotion; ability to learn a procedure; ability to react to a changing situation; ability to perform during emotional stress.

Discrimination of Visual Motion and Locomotion

The motion picture, as has already been explained, can represent not only motion of an object but also locomotion of the subject or observer. The sensory basis for this fact is complex and is discussed fully in chapter 9. The peculiar type of motion stimulation which gives rise to motion of the *subject* instead of motion of *objects* is there termed *retinal motion perspective.* The phenomenon of "motion parallax" as a cue for distance perception is a special case of it—the special case in which the line of regard is at right angles to the line of motion.

This capacity to represent motion and locomotion makes it theoretically possible for motion picture tests or experiments to measure the ability to judge any *visual* aspects of locomotion which may have to be judged in practice, such as velocity, direction of one's movement, or momentary position with respect to reference points. Tests measuring such ability are of particular significance in a program of aviation testing and research.

Judgments of a number of different kinds of visual movement may be obtained with motion picture tests. Three different kinds of judgments all involving the perception of velocity are obtained

with the existing Speed Estimation Tests CP205B–I, CP205B–II, and CP205B–III. The intercorrelations of these three tests are very low (from .02 to .15), despite the fact that the types of performance demanded by all three seem, on the surface, to have a good deal in common; there is no unitary psychological function relating them. The fact that it can be given a name is misleading. The judging of moving objects is evidently a complex matter. Different sorts of measurements are obtained in the Minimal Movement Test CP213C and the Drift Direction Test CP221B, in which, respectively, judgments of the existence (i. e., absolute threshold) of motion, and of the direction or "drift" of a visible motion alongside of a parallel line, are required. These types of movement discrimination are similar to those required in the job of synchronizing a bombsight. The correlation between them is also low, about .20. All the above facts are described in chapter 5.

A test could be devised in which an object (a spot or model plane) moves across the screen at a rate which is constant except when it passes a specific point, at which instant its speed would be slightly increased or decreased. The change in speed would occur at one of five points, labeled A, B, C, D, and E. The task of the subject would be to indicate which of these letters marks the point at which the change in velocity took place. Such a test would not depend upon judgments of coincidence as indirect measures of velocity discrimination, as do Speed Estimation Tests CP205B–I and CP205B–II, nor would it be as highly dependent upon memory, as is the case with Speed Estimation CP205B–II. Thus it might provide a simple, more direct, and possibly more useful, measure of velocity discrimination. It may be noted that this method of measuring different sensitivity to velocity is practicable mainly because of the use of animation photography, which simplifies the problem of presenting and controlling complex moving stimuli.

Another aspect of motion, the estimation of which could be measured by means of a motion picture test, is acceleration or deceleration. The test could depict a model plane moving across the screen. The plane would move with a constant deceleration for a certain distance, and then disappear. The task of the subject would be to estimate at which of five points the motion would have stopped if it had been continued. The importance of this type of judgment for the pilot's or bombardier's task is not known. It is not expected that the function measured by such a test would be highly related to judgments required in the other speed estimation tests.

In comparison with the motion discrimination tests described above, and with the three existing Speed Estimation Tests, a test for estimation of subjective velocity would be interesting. It

703313—47—3

should be included in any systematic study of ability to judge motion and speed, since it might prove to be unique in relation to the other tests in the field and its validity for pilot selection might be higher. The simplest way to construct such a test, short of special effects photography, would be to photograph the view ahead from a vehicle moving along a straight road without any regularly recurring features such as telegraph poles. Shots would be made at a number of just noticeably different speeds, and enough duplicate shots of the same speed with different terrain would be provided to yield a sufficient number of 6-second items for the test. Discrimination would then be independent of the terrain photographed. Judgments could be based on successive comparisons, or, more advantageously, in terms of absolute judgments of miles per hour after showing and "teaching" a series of stated velocities.

The capacity of motion pictures to represent locomotion, direction of one's movement, or momentary position with respect to reference points is illustrated by the following examples: The Landing Judgment Test CP505E measures ability to judge the direction of locomotion in a landing glide by identifying the spot on the runway toward which the glide is aimed. The Landing Orientation Test CP106A measures ability to judge one's momentary position in a course circling an airfield, i. e., one's location during flight in a traffic pattern. The Altitude Judgment Test, described in the next section, is intended to measure judgment of another type of momentary position, namely altitude above the ground during a landing glide. A test is possible measuring the ability to judge these various visual aspects of locomotion in combination, to weigh them, and to make a prompt decision as to action. The scenes presented would be what the trainee sees in critical flying situations.

Perception of Space and Distance Particularly During Flight

A test of distance estimation employing still photographs already exists in experimental form. This test is described in chapter 9 of this report. One of the important cues to the judgment of distance is motion. One form of this cue is that termed "motion perspective." Distance can be represented in still photographs which do not include motion, with a surprising degree of adequacy, but it is reasonable to suppose that a still truer representation of the situations in which distance judgments are required would be attained by a test in which the motion cue is present. One possibility of test development in this field is as follows: The same scenes that are used in the present still-photographic test can be taken with a motion picture camera moving at a right angle to the line of regard. The same type of judgment would be required as in

the case of the printed test. This test, in motion picture form, with the addition of relative motion, should have the effect of making the relative distances in the scene more determinate and, hence, making the test more valid as a measure of the ability to estimate distance in real situations.

A test of the ability to estimate altitude could be devised which depicts the approach to a landing strip during a straight glide, in which the subject is required to match the altitude at some particular point of the approach with the same altitude chosen from five alternatives, subsequently given. In this case the alternatives would be presented sequentially as time intervals during the glide by having a voice announce them during the repeated scene. A test of this sort would depend to some extent upon memory. It would include motion as one of the important cues to distance estimation, and also would resemble the real situation in which distance discriminations are required, by utilizing judgments made in terms of a temporal sequence.

Maintaining Orientation During Locomotion

A major difficulty encountered in the design of tests of orientation is the tendency for the examinee to translate the task into the terms of a numerical or verbal schema, such as the points of a compass. This tendency probably accounts for the moderately high correlations often found between orientation tests and tests of numerical or verbal ability. One method of reducing the probability of this undesired intellectualization is to avoid presenting the orientation problem in terms which suggest compasses and other devices, insofar as possible. This, however, seldom solves the problem completely. Another method is to present the spatial situation in such a realistic way that doing it intellectually is much more difficult than "feeling oneself into it." Presumably the type of orientation ability which is displayed under the latter condition is the type for which measurement is desired. In the Flying Orientation Test CP107A described below, this method has been employed.

On the screen is shown a view of terrain as seen through the open bomb bay doors of a plane in flight. The course which the "plane" follows consists of several short legs of equal length connected by turns of 90°. The examinee must imagine that he is flying in the plane, and keep in mind the direction of the starting point. At the end of each flight, the subject responds by indicating, by reference to a lettered circle shown on the screen, the direction of the starting point from his present position. In experimental administrations of this type of item, the subjects found it impracticable to keep track of the direction of the starting point by means of implicit verbal symbols such as "right," "left,"

"north," "south," etc. This was apparently due to the fact that the direction (of the starting point) changes not only at a turn, but continuously during certain of the legs of the course. For example, after one turn the starting point may be to the right, but as the plane continues to move along the next leg, the starting point drops farther and farther behind. The symbolic process for the representation of these continuous changes would have to be relatively complex, and the subjects are in a sense "forced" to adopt the simpler expedient of imagining themselves moving in space. It is believed, therefore, that the capacity for representing subjective motion, present in motion pictures, has made possible a unique solution of an important problem in the construction of orientation tests.

Another possibility exists, using film from the above test. It is possible to require the subject to visualize the path through which an airplane has flown, after having watched the movement of the ground underneath, during each flight. For each item, the screen would depict the ground as seen through open bomb bay doors during a flight involving several turns. At the end of the flight, the subject would attempt to identify the path from five alternatives on the screen. This test strongly resembles the Flying Orientation Test described above, except that a different type of response is required. The same items could, however, be used.

Ability to Learn a Procedure

The characteristic of motion pictures which makes possible the representation of a series of acts in their natural sequence and tempo provides the opportunity of testing for the ability to learn a specified procedure, not by verbally memorizing a series of names or written descriptions but by actually visualizing the series of acts as they naturally occur.

A test could be designed to measure memory for a series of discrete events or acts. Several types of aircrew duties seem to require this function, particularly bombardiering, in which the student must learn to perform a rather long series of more or less disparate activities in a specific order. In abstract form, such a test might be constructed as follows: a sequence of responses (i. e., a procedure) would be presented visually on the screen and also orally by means of the sound track, to be followed by another sequence showing an individual following the procedure (e. g., adjusting knobs in a certain order) with certain errors. The errors in the visually presented procedure would be identified by letters and recorded on a standard answer sheet. With the addition of realism, such a test might become a work-sample aptitude test.

The Successive Perception Tests CP509C–I and CP509C–II

measure visualization of events or comprehension of temporal patterns by requiring the examinee to visualize (a) a total pattern exposed a small part at a time in continuous sequence, and (b) the total path of movement of a continuously moving spot. Other tests of visualization are possible, all involving the *sequence* which is provided by motion pictures.

Ability to React to a Changing Situation

The cinematic characteristics of motion, sequence, and pacing make it possible to present either a single changing situation, i. e., a situation "flowing" or proceeding in time, or a number of such situations occurring simultaneously on the screen.

The Flexibility of Attention Test CP411E and the Integration of Attention Test CP415A, attempt to measure the ability of an individual to observe a number of different activities at the same time, by presenting on the screen five continuously moving dial indicators. The task is a speeded version of a similar real job, and the attempt is made to represent the natural irregular pacing of the real situation under which "division of attention" is required.

It may be that the task can be better described as one in which a number of continuing *intentions to react* must be simultaneously kept up, while at the same time observing a number of irregularly changing stimuli, such as dial readings to which the appropriate reactions must be made. This description of the task emphasizes the factor of keeping a number of things in mind at one time, while remaining alert to observe signals of danger. A test could be designed to represent this latter type of task. There are a number of possible ways in which such a test could be presented in motion picture form, involving motion in clocks, dials, etc., shown on the screen, along with a task which requires the examinee to keep in mind several discrete activities. One simple form would be as follows:

The subject is given the two *continuing* tasks of recording the readings of dials A and C every third minute (which is indicated by a clock shown on the screen) and the readings of Dials B and D every second minute. These dials, A, B, C, and D, are graduated in units which run from 1 to 10, and their readings vary continuously, though relatively slowly, throughout the test. The clock from which time is read has a minute hand which rotates once in 30 minutes, and a second hand which rotates once in a minute. Thus its movement must be followed, not continuously, but with a reasonable degree of care in order for reactions at the critical movements to be made correctly. The task described could be made more or less complex, but it meets the specifications of providing the examinee with *a number of self-maintained sets to make particular reactions at specific times.* In addition to these

basic activities which the subject must himself initiate at definite intervals, he may also be required to record the "going wrong" of certain other dials which are numbered 11, 12, 13, 14, and 15, whose movements are irregular and relatively rapid. The test could be so constructed that any desired proportion of the items required reactions at specific intervals and at irregular intervals. It is believed that a task of this sort would permit the measurement of the ability to maintain attention to a set of regularly recurring, possibly monotonous, activities in the face of other demanding and distracting tasks.

Another interesting possibility exists of changing the usual task in printed mechanical comprehension tests by presenting similar material in motion picture form. A common type of item in printed tests depicts a series of gears or pulleys with direction of movement of the driving gear (pulley) indicated by an arrow. Usually, in order to answer the item, the examinee must *visualize* the movement of the driver and infer the movement of each subordinate part of the system. His ability to imagine such movement seems, in fact, to constitute in large part the function being measured, which has a demonstrated usefulness in the aircrew classification battery. But this ability to visualize movement from static diagrams does not constitute the comprehension of mechanical principles or at least of dynamics. It is conceivable that a motion picture test of mechanical comprehension might measure a quite different function. In this case, the motion, rather than being imagined, would be presented directly on the screen. A simple item might take this form: On the screen are shown two unconnected gears of equal size. One is rotating clockwise at a regular rate, the other rotates counter-clockwise at half the rate of the first. In a printed test booklet are shown five possible connections of these two gears, with auxiliary gears, and the subject must choose the one which could correctly reproduce the movement shown on the screen.

It seems evident that such a test would measure understanding of mechanical principles with particular reference to *dynamics,* a function which might have considerable validity for predicting the success of the bombardier or flight engineer trainee. Although the degree of relationship between the motion picture version and the various printed versions of mechanical comprehension tests is difficult to predict, the task presented by the motion picture test seems a unique one.

Ability to Perform Under Emotional Stress

Motion picture tests of the efficiency of performance under conditions of stress or emotion are of two general types: that in which the stress is induced by the difficulty of the task and the consequent sense of failure and inferiority (e. g., by the pacing of

the test), and that in which the stress is induced by the portrayal of events and experiences arousing fear, shock, anger and the like, (i. e., by the realism of the test). One test of each type was planned and studied in a preliminary way.

The Reaction to Stress Test is an adaptation of the Plane Formation Test CP805C, in which both the exposure interval and the response interval are reduced progressively until it becomes impossible to see and record the complete pattern constituting the item. Five marks must be made in order to complete each pattern successfully (i. e., subjectively); the test is scored, however, on the basis of each mark taken separately. Interspersed with these items at irregular intervals are a number of relatively "easy" items, in which both the perceptual task and the recording of answers can be accomplished within the time allowed. The scores made on these particular items would reveal whether or not the subject has been able to persevere at the task in the face of discouragement or stress arising from failure. Only these items are scored, although the examinees are not aware of this fact. The possibility of constructing this test obviously depends upon the characteristic of the motion picture medium which permits the pacing, in this case the speeding up, of both the exposure and the response intervals.

As another possibility, the Observation Test (Susceptibility to Emotional Stress) was designed to measure an individual's ability to control the degree to which he is influenced by emotion-inducing situations. The scenes should be chosen both for their realism, and for the extent to which they induce in the subject the experience of taking part in the events shown. A scene showing an American soldier in brutal hand-to-hand combat with an enemy soldier, with the camera aimed in such a way that it has the "point of view" of the American soldier, is an example of the type of scene which seems useful for this purpose. The task required of the subject would be to remember and observe *certain minor and irrelevant events* which occur on the screen and which are a part of the total scene presented but not of the highly emotional "story" told by the scene. Presumably the extent to which the performance of the purely perceptual task is interfered with would yield a measure of susceptibility to emotional stress. The noticing of the minor fleeting events would require "cool observation." As the discussion in the earlier part of this chapter indicates, the advantages of this design are believed to reside in the fact that it is difficult for the examinee to dissociate his experience of the emotional scene from the observational task itself, since he cannot look away from the screen, and since the elements to be observed and remembered form a part of the scene itself.

CHAPTER THREE

Technique of Construction of Motion Picture Tests*

PRELIMINARY CONSIDERATIONS

Motion pictures can be presented either in a theater equipped with dual 35-mm. sound projectors run by a trained operator, or in a smaller room, of classroom size, with a single 16-mm. sound projector. The latter procedure is much simpler. Sixteen millimeter sound projectors can be operated by anyone with a moderate degree of training, and they are mechanically reliable. By observing a number of fairly simple precautions, 16 mm. films can be shown with a degree of success approximating commercial films and far superior to that of ordinary "home movies."

Army training schools made use of this fact on a wide scale. All training installations were equipped with classrooms set up for 16-mm. sound projection and a great variety of films for instructional and other purposes was shown. The existence of such equipment at all centers where psychological testing was being carried out in the AAF was the basic fact which made it possible to plan an experimental program of motion picture tests.

It was clear from the outset, therefore, that psychological test films should be of the 16-mm., type and that the directions for the test should be incorporated into the film itself whenever possible. It was estimated that from 6 to 20 prints would be required for each test to supply the different testing agencies, and possibly more if the test should prove valuable at a later date; consequently the film had to be capable of being duplicated. It was also assumed that a motion picture test should ordinarily not exceed 30 minutes in length and should if possible be kept to a length of 15-20 minutes; it was therefore possible to produce all test films on a single reel and to obviate the necessity of changing reels during administration.

Sixteen millimeter motion picture film can be produced in three general ways. It can, as a first possibility, be photographed, processed, edited and projected on 35-mm. film, and then transformed

*Sections of this chapter were drafted by Russell W. Bornemeier, Ralph M. Eisenberg, and Gerald M. Slater.

into 16-mm. film by "reduction printing." This permits the technical processes of film making to be carried out at the level of commercial studio work, and it is of course superior to that of other methods. This was the method usually employed in producing training and indoctrination films. As a second possibility, the photographic and other work on test films can be carried out with 16-mm. film, and prints made directly from a completed 16-mm. negative. Since 16-mm. photography is within the range of accomplishment of a semitrained amateur like the psychologist, this has advantages for the producing of experimental films such as tests, where the experimenter needs to try various possibilities and control the making of the film himself without having to work through technicians. The sound has to be added with professional studio equipment. Nevertheless, certain types of motion picture photography useful in testing are beyond the capacity of this method, such as animation and various special effects. At present, these accomplishments are seldom possible except in studios with 35-mm. equipment. The third possibility is strictly amateur film production, employing 16-mm. "reversal" film. Its advantage is that no laboratory facilities are required, since the film is processed and returned without extra charge by the manufacturer, a positive image being produced directly on the same film that was exposed. Such film cannot take a sound track and it cannot be duplicated. Additional positive prints can only be made by means of an intermediate negative, the effect of which is to reduce the photographic quality of the product. The method is therefore inappropriate for anything but personal films of limited interest, or wholly preliminary experimental work.

Of these three methods of producing 16-mm. films, the first was the one principally employed in the program of experimental motion picture tests to be described. One test only was constructed by the second method with 16-mm. film throughout all stages, being originally photographed with the semiamateur equipment and skills available to a research unit as distinct from a professional studio. The first method is wholly dependent on the collaboration of a studio operating on commercial standards. The second method would have been advantageous and practicable for a number of tests, and the plans made by the Psychological Test Film Unit called for some of this type of work. They failed of application because of the critical shortage of motion picture cameras which persisted for some time after the latter part of 1943 when the unit was established.

The stages in the production of a psychological test film are usually as follows:

1. Preliminary tryout of the problem, task, or performance to be incorporated into the test as items or trials. This can be done

by preliminary 16-mm. photography or by various makeshift methods such as models moved by hand, and it does not involve the services of motion picture technicians.

2. The preparation of a script for the introduction to the test, giving both the shots to be photographed and the text of the accompanying directions for the test to be recorded on the sound track. The visual material, including subtitles if any, must be roughly synchronized with the voice. Along with the script goes a set of *specifications* for the photographing of the various shots which will make up the items of the test and for the order in which they will be spliced. This stage also is performed by the psychologist.

3. The filming of the specified scenes. This requires a motion picture studio for some types of photography or at least a skilled photographer for others.

4. The processing of the film in a laboratory and the making of prints; the cutting and splicing of this film in the form of a "work-print" which will synchronize with the script. The latter work is performed by a film editor or cutter who must work in close cooperation with the test constructor.

5. Tryout of the test in this "work-print" form for purposes of item analysis and revision of the script and specifications. Since the film is silent, the directions for the test are read aloud by a test administrator as the film is projected before a group of candidates.

6. Revision of the test film. This will require new photography only if the test specifications were ill-conceived. Usually revision can be accomplished by eliminating the ineffective items and by re-editing the work print. Special effects photography may be added at this point.

7. Recording of the revised script on a sound track, final processing and editing of the film and the manufacture of final prints for distribution.

It will be seen that motion picture tests are somewhat cumbersome to revise. This difficulty serves to emphasize the need for careful preliminary analysis and tryout experimentation before specifications for the final form of the test are prepared. Experimentation should include statistical analysis. In this way final decisions can be arrived at concerning item difficulty, serial order of items, and length of test, and hence the possibility of a necessary revision following the completion of the test is reduced.

In the discussion of the technique of construction of motion picture tests to follow, the assumption is made that at least the final production of the film will be accomplished by technicians. The skill and facilities necessary for a finished and professional-

appearing motion picture test are at present only available commercially.

CONSTRUCTION OF TEST ITEMS

Answering and Scoring Procedure

With one exception, all motion pictu tests constructed by the Psychological Test Film Unit were designed to have responses made on standard or modified I.B.M. answer sheets scorable by machine. This requirement does not impose a serious limitation upon the possible nature of the test items, but it may necessitate considerable ingenuity at times in item construction. Basically, of course, answering procedure with the motion picture medium introduces no peculiar problems except that it is necessary to regulate the duration of the recording interval in such a way that the slowest examinee has adequate opportunity to record his answer. For ordinary tests this interval has been experimentally determined to be from 5 to 8 seconds. It may, of course, be desirable to speed up the recording interval in a test which measures the ability to perform some intellectual or perceptual task in a limited time, e. g., to induce a feeling of stress. The requirement of machine scoring makes necessary serious consideration concerning the exact nature of the answering procedure before item construction begins.

In general, answering procedures employed in the case of printed tests are also satisfactory for motion picture tests. These procedures include the use of standard I.B.M. multiple-choice answer sheets involving 2, 3, 5 or 15 alternatives, in which the alternatives may be either lettered or numbered. A novel answering procedure must sometimes be developed, however, for the novel psychological functions which motion pictures are capable of testing.

Answer sheets can be adapted to special uses by overprinting in such a manner that routine machine scoring is unaffected. The spaces on the answer sheet in which the examinee marks his answers may be altered, by special printing, into special spatial arrangements or may be given designations particularly appropriate for a given test. An example of an answer sheet in which the answer spaces are arranged in a special spatial pattern is that used for Plane Formation Test CP805C. In this test, the examinee is required to fill in spaces on the answer sheet corresponding to the sections of a grid appearing on the screen; the grid contains 25 squares upon which appear five small airplane silhouettes. The examinee is required to reproduce the pattern made by the five planes. An illustration is given in Chapter 5. It is interesting to note, therefore, that by proper arrangement of the answer spaces, an answer sheet can be used for recording judgments of

spatial patterns and extents. By designation of answer spaces as U, D, and S for Up, Down and Straight, respectively, instead of A, B, and C or 1, 2, and 3, it is possible to obtain judgments of the direction and extent of movement as in the Minimal Movement and Drift Direction Tests. This is also illustrated in chapter 5.

For certain types of tests it is desirable that the examinee view the screen for relatively long periods of time without, interruption for the clerical task of recording his observations. A transcription technique has been developed for this answering situation. The examinee is given a work sheet on which he marks his answers continuously as events occur on the screen without shifting his attention from the screen. Upon the completion of the test the examinee transcribes his answers from the work sheet to a standard answer sheet.

Techniques for Construction of Preliminary Test Items

There are a few general techniques for the construction of preliminary test items. The appropriateness of any specific technique for any given test will depend upon the nature of the test and other considerations such as equipment and facilities available.

It is to be understood that the product of these techniques is not to be used as a final form of the test. These techniques yield a preliminary form, without instructions recorded on the sound track, which can be administered informally to small groups of subjects. Preliminary administration will give information concerning item difficulty, item exposure time, recording interval, and serial order of items. By this analysis information can be obtained for the preparation of exact specifications for the final form of the test.

When preliminary test construction begins, one of two general situations usually prevails. Sometimes films are available which, by editing, can be made into a form satisfactory for tryout purposes; more generally a form of the test must be photographed, sometimes on 16 mm. film. It may be possible to construct and try out items in the absence of photography by the use of mock-ups or other mechanical devices. Sometimes a combination of the above procedures is necessary.

Production by Editing of Existing Motion Pictures. The production of the Motion Picture Comprehension Test CI625A is an example of the use of this procedure. A large number of motion pictures were reviewed in search of sections suitable for the construction of this test. Prints of the suitable sections containing satisfactory material in the appropriate length were obtained. Test instructions and preliminary practice items were written and test questions on the material in the sections were prepared. These sections were spliced together and this form of the test adminis-

tered with a narrator substituting for the sound track. From these preliminary administrations information was gained concerning the nature of the test so that specifications for the final form of the test could be written.

Production by New Photography. Usually it is impossible to find existing motion pictures from which to construct items, in which case it is necessary first to devise the items and then to photograph them. The material to be photographed for an item might be a carefully rehearsed scene (e. g., Observation Test CE307A); or some type of movement of a drawing, map, or photograph (e. g., Flexibility of Attention Test CP411E). The essential feature of the procedure is preliminary photography of the items so that they can be spliced together and administered for statistical analysis.

The production of the Flexibility of Attention Test CP411E is an example of this kind of test construction. A panel was constructed on which appeared five schematic instrument dials having simple indicators. The test items were obtained by photographing the panel and dials while operators, standing behind the panel, moved the indicators in a predetermined manner. These shots were spliced together, the instructions were provided vocally, and the test was administered in this form to groups of subjects so that a statistical analysis could be performed. Following the analysis, the test was modified wherever necessary, specifications written, and the test constructed under supervision in an army motion picture studio.

It is possible in certain instances to construct and try out items by the use of mock-ups or models. In this case preliminary photography becomes unnecessary. An example of motion picture test constructed in this manner is the Successive Perception Test I (Moving Slot) CP509C–I. In this test, a slot in an opaque screen moves over a black geometrical pattern exposing it successively from top to bottom. The patterns vary in nature and complexity from item to item. The patterns were drawn on cards and exposed manually with a slot of the sort described. In this way the test was tried out in preliminary form with groups of subjects. Statistical analyses were performed and, finally, specifications were drawn up for animated motion picture photography at a studio.

Other devices are useful in the construction of test items where preliminary photography is not indicated. In addition to mock-ups and models there are ordinary photographs, slide projection of pictures, and drawings. Full use of such devices can effect substantial savings in time and expense.

Production by a Combination of New Photography and Extensive Editing. In some cases the most effective and convenient way to work is to draw up specifications for a basic test situation. By

means of these specifications, a production studio is able to provide film in 35 mm. form from which a series of test items can be obtained by selecting and editing. By appropriate cutting and arrangement of the items along with a narration to be read as the film is projected, a tryout form of the test can be produced. This form, then, is administered to groups of subjects and with further revision, final specifications for editing and sound can be written.

An example of this sort of test construction procedure is the Landing Judgment Test CP505E. The rationale and the basic test situation for this test were carefully elaborated and specified. Because of the difficulty of predicting how and with what success subjects would make the judgments desired, it was necessary to photograph the test situation under a number of varied but closely determined conditions. By experimental administration, selection, and modification of these basic shots, it was possible to determine the length and difficulty of the final items, to decide upon their order and arrangement, and to write the instructions. The tentative experimental form of the test could be modified and revised, until a satisfactory form was evolved. Specifications were only then written for special photographic effects, and for sound.

The preceding sections have dealt with various techniques for producing a tryout form of the test. The following section is concerned with administration of this form of the test.

ADMINISTRATION OF PRELIMINARY FORMS

Administration to Sophisticated Subjects

It was the experience of the Film Unit that the assistance of an audience of psychologically sophisticated observers, or "test-wise" subjects, is extremely valuable at the preliminary stage of test development. Such subjects can provide introspective accounts of the psychological functions involved in the task and professional opinions of the success with which the test measures these functions. They can also note unwanted cues to the correct answers which have escaped the test constructor because of his familiarity with the test. The effectiveness of this procedure was illustrated during the construction of the Flying Orientation Test CP107A.

It was desired in this test to measure the ability to maintain directional orientation when flying, together with the allied ability to visualize a flight path already flown. In the original form of this test the subject was required to state the compass direction in which he was traveling at the completion of the item, *i. e.*, a "flight" shown on the screen. Preliminary administration of the test to a critical audience established the fact that the desired measurement of the ability to maintain directional orientation was

not obtained. Instead the task set by the test could be performed in an intellectual manner by the use of verbal memory. Comments by the audience suggested modifications of the test which prohibited intellectual solution. In the final form, the task of the subject is to maintain his orientation by watching the ground as it wheels below him, so that at the end of a flight involving several turns he can indicate the direction of the starting point from his present position. It is believed that this form of the test successfully measures the function for which it was designed.

Preliminary administration of the test to psychologically sophisticated subjects can also provide information concerning the technical aspects of the test's construction. The adequacy and clarity of the test instructions can be criticised. Estimates can be secured as to the appropriateness of the item exposure time and the length of the recording interval. These intervals can be crucial variables in the construction of certain tests. Such an expert group also can note any disadvantageous arrangement of the serial order of items.

Administration to Unsophisticated Subjects

When the test is developed as far as possible by the foregoing techniques, it is necessary to administer it to unsophisticated subjects for purposes of statistical analysis. It is desirable, of courses, to administer the test to an actual sample of the population for which it is designed.

Statistical analysis at this stage of development usually consists of item analysis and the study of item difficulty, test reliability, sometimes intercorrelations with other existing tests, and information concerning the effect of practice on the test. In certain cases, studies may be made on the variable of seating arrangement and illumination, and in some cases preliminary validities may be secured.

By the use of the foregoing procedures and preliminary test administrations it is possible to write the final script and specifications for the test. The following two sections deal with their preparation.

PREPARATION OF THE SCRIPT FOR THE INTRODUCTION TO THE TEST

The "script" represents the film cutter or editor's guide for the voice, the titles, and the portrayed events of the motion picture test in its final form. It represents what is called the "continuity" of the film. As written by this unit, the part of the script to be "narrated" appears on the right-hand side of the page, and the accompanying action is given on the left-hand side of the page. This arrangement enables the cameraman to get an overall idea of

the nature of the test as well as serving the film editor as a working guide in cutting and assembling the film in its final form.

The part of the script describing the "action," or visible aspect of the film, should enable the cutter to determine the proper sequence of shots and their duration. The details for the photography of the action that appear in the script are submitted as a set of specifications and will be discussed in the next section. The script for a motion picture test is similar to an ordinary "shooting script" for a motion picture of the entertainment or educational type, in that, when read, it should make clear the content and nature of the film which will be projected. Its emphasis, however, is particularly on the introductory portion of the test which contains the directions and sample items. The main body of items for the test does not require a script. The "shots" or successive units of film which constitute the items of the test can best be described in a set of specifications primarily for the use of the cameraman. A portion of the script written and submitted for the production of the Flying Orientation Test CP107A is presented below as an example of the form used by this unit. This script was prepared for 35-mm. studio photography. The test which was actually produced, however, had to be made with 16 mm. film and only semi-professional equipment, and was based on a simplified script.

SCRIPT FOR FLYING ORIENTATION TEST

Main Title:	Aviation Cadet Testing Program (Air Corps Symbol)
Cut To:	Flying Orientation Test
Smaller Letters	Psychological Test Film Unit Santa Ana Army Air Base December 1944
Title accompanying voice:	This is a test of your ability to remain oriented while flying a complicated flight path.
Cut to view of moving terrain seen through bomb bay doors.	What you see on the screen represents what you would see if you looked down through open bomb bay doors from a plane in flight. Notice how the ground appears to move as you fly over it.
At word "this," terrain begins a turn simulating plane turning 90° to the right.	This is the way the ground appears as you make a 90° turn to the right. (Pause)
Plane completes turn and flies straight again.	

At word "this" plane begins a 90° turn to the left.	This is how it looks when you turn 90° to the left. Remember your *plane* is turning to the left.
Plane completes turn and flies straight again.	In this test all the turns will be 90° turns and the paths between turns will always be the same length.
Cut to title accompanying voice.	Your task is to remain oriented at all times during a maneuver so that when the flight ends you can visualize the path you have flown and know the direction you would have to go to return to the starting point.
Another title accompanying voice.	Watch this practice flight. The path the plane will fly has been marked out in advance with dotted lines.
Cut to moving terrain through open bomb bay doors. Path is marked out with dotted lines on photograph. Circle super-imposed over terrain faintly with heavier arrow pointing back.	Starting here and flying straight, your starting point is now behind you. Note the circle with the arrow pointing back toward the start. This arrow, which will always point toward the start, is to show you how you should visualize the direction of your start during the test.
Plane turns to the right. Arrow swings with the terrain to point toward start at the right.	Your plane is turning to the right. The arrow now points to the right toward the start.
Plane flies straight. Arrow moves over terrain, swinging slowly to point down, following direction of start.	As you fly straight, your start is still to the right, but it is getting further and further behind you.
Plane stops. Circle and arrow fade to circle with lettered arrows, with large arrow at "D."	As the plane stops, the arrow is pointing toward your start, which is to your right and behind you.
Camera view expands to show all of terrain, with path flown marked with dotted lines. Small planes along path, in circle with arrow pointing toward start.	Here is an overall view of the path you traveled and the way in which you must visualize it. Your task is to choose the lettered arrow that points toward the start.
	Imagine yourself seated in the plane, which is always flying in direction "A." Your start is to your right and behind you, in direction "D." Remember, this overall view is only to show you the path you have flown. During the test you must think of the direction of the start from your position in the plane.
Cut to close-up of answer sheet. Hand marks space "D" of Item No. 1.	Mark the space under "D" as the answer to question number one.

703313—47—4

Cut to Item No. 2 for 1 sec.	**Ready for practice flight number two.**
Cut to moving terrain through bomb bay doors. Arrow in circle pointing down toward start. Path marked out with dotted lines.	**As you fly straight, your start is behind you.**
Plane turns right. Arrow turns with terrain, pointing toward start.	**You are turning to the right, so the arrow points toward the start at the right.**
Plane flies straight. Arrow moves over terrain, swinging down to keep pointing at the start.	**Your start is still to your right but it is somewhat behind you.**
Plane turns right. Arrow keeps pointing toward start.	**You are turning to the right again. Your start is now in front of you and to the right.**
Plane flies straight and stops.	**Your start is now directly to your right.**
Camera view expands to show all of the terrain with path flown marked with dotted lines. Small planes along path inside a circle with an arrow pointing toward the start. Circle and arrow showing the direction of start during flight fade to circle with lettered arrows. Large arrow at "C."	**Here is an overall view of your plane and the path you flew. Note that the circle around your plane has lettered arrows showing directions from your plane. Thus arrow "A" always points ahead of you, in the direction you are flying. The arrow which showed the direction of the start during the flight is pointing in direction "C." This arrow and the dotted lines will be shown only during the practice flights to help you visualize your path and the direction of the start. Mark the space under "C" as the answer to question number two.**

The script continues in the same form through five practice items, and the instructions end with the following statement:

"We will now begin the test. Remember, you must visualize your flight and know the direction you must go to *return* to the start."

After the instructions are completed, the test begins with Item 6 and continues through Item 50. Five seconds of answer time are available after each item.

PREPARATION OF SPECIFICATIONS FOR PHOTOGRAPHY OF THE TEST

Specifications should include all the technical data necessary for photography of the test. The content will necessarily vary with each test, depending on the type of photography, e. g., aerial photography, model photography, or animation, but all specifications should include a description of the basic scene to be photographed, a description of the exact shots for each trial including the practice trials, and a list of titles.

It must be kept in mind that the cameraman gets all his information from the specifications and script submitted to him. It is therefore necessary to submit precise written explanations, diagrams and sketches so that he will be able to photograph the necessary film with as little personal supervision as possible from the psychologist.

As an example of specifications prepared and submitted by this unit, a section of those for the photography of the Flying Orientation Test CP107A as originally conceived is presented below.

SPECIFICATIONS FOR PHOTOGRAPHY OF FLYING ORIENTATION TEST

(35 mm. special effects photography)

1. *Description of Scenes.* The test is made up (primarily) of 85 shots which will be called "flights" including five practice flights. These are preceded by instructions and one special flight used as an example.

a. *Description of a "Flight."* A "flight" should simulate the appearance of the ground as an observer would see it looking straight down through the open bomb bay doors of a plane. An aerial photograph is substituted for the ground in the filming of the test; the axis of the motion picture camera should always be perpendicular to the photograph. In each flight the airplane appears to fly a path made up of several straight legs of constant length. These paths are drawn accurately in the drawings which accompany this description. In these drawings, a frame size of 1 inch by 1.33 inch is assumed. The starting and stopping points are drawn so as to be in the *center* of this frame. The 90° turns are made without bank, the axis of the camera always remaining vertical.

b. *Bomb Bay Doors.* Extending out from the right and left edges of the frame are representations of open bomb bay doors. These may be either photographs of real bomb bay doors superimposed along the inside edges of the frame or they may be painted likenesses. The appearance of the screen, with bomb bay doors, is shown during a typical flight in figure I. The open space in the center of this frame should consequently measure 0.95 inch by 1.0 inch on the scale of the photograph. [This illustration is shown in chapter 5.]

c. *Terrain Photograph.* For the purposes of writing specifications it is assumed that the terrain shown between bomb bay doors will be a portion of the aerial photograph which accompanies this description, taken from an altitude of 20,000 feet. If it is desirable to enlarge this photograph, the dimensions and distances, in terms of real space on the ground should remain exactly the same.

d. *Specifications for a "Flight."* On the 1 inch by 1.33 inch frame described above, the flight path should be photographed so that the center of the frame will trace a path over the photograph having the following characteristics: The length of a leg from the finish of one turn to the beginning of the next is 1.4 inch. The 90° turns are to trace a path which corresponds to an arc of 90° having a radius of 0.9 inch.

The flights will be photographed at the following speeds, using the 1 inch by 1.33 inch frame:

Speed 1—3.5 secs. per inch (Example flight, Flight 1).

Speed 2—3.0 secs. per inch (Flights 2 through 5).
Speed 3—2.1 secs. per inch (6 through 30).
Speed 4—1.7 secs. per inch (31 through 59).
Speed 5—1.3 secs. per inch (60 through 85).

The permissable variation in speed is 5 percent. The speed on the turns should be exactly the same as that on the legs with the axis of the camera remaining vertical always. The exact number of seconds required for each flight is given on the drawing of each flight.

e. *Circle and Arrow.* During the first 3 practice flights and following flights 4 and 5, a circle and arrow are superimposed over the terrain in the center of the frame. The appearance of the circle and arrow is indicated in figure II. The outside diameter of the circle (in the 1 inch by 1.33 inch frame) is 0.25 inch and the length of the arrow is 0.19 inch. The circle and arrow should be white and transparent, The circle should be clearly visible, but sufficiently transparent so that the ground can be seen through it. The arrow should be of slightly greater density, giving more contrast with the ground below.

The circle and arrow move "with the plane" over the ground, remaining in the center of the opening between the bomb bay doors. The circle does not rotate, but the arrow can turn through 360°, pivoting at the center of the circle.

During a "flight" the arrow must turn so that it always points toward the place on the photograph over which the "flight" started, i. e., the starting point. As the "plane" makes a turn to the right, the arrow rotates around approximately 90°. As the "plane" continues along the second leg, the arrow drifts slowly backward to 135°, following the starting point.

Except when the arrow is pointing at 0°, (straight athead) or 180° (straight behind) it will be moving continuously during a flight.

f. *Broken-Line Path.* In the first three practice trials, the path the "plane" flies is marked "on the ground" with a broken line. This line is made up of white opaque dashes and appears to issue from the pivot of the arrow in the center of the circle. It appears to trace on the ground the path which the plane flies in the air.

When the arrow described above is pointing to 180° it will coincide with the direction of this broken line. The broken line, however, should always be visible through the arrow. The appearance should be given f the circle and arrow up in the air above the broken line which is tracing the path on the ground.

g. *Scene Showing Entire Flight Path.* The complete path of the plane on the ground is shown at the end of each of the five practice items. After the "plane" stops, the camera recedes from the ground, still centered on the end point of the path, until the whole path flown and the present location of the starting point are visible on the screen. *The stopping point of the flight should retain its orientation and remain in the center of the screen.*

When the camera recedes, only the ground and the broken-line path appear to get farther away. The circle and arrow as well as the bomb bay doors recede with the camera, i. e., they retain their original size.

h. *Circle with Eight Lettered Arrows.* The circle with the 8 lettered arrows is cut in at the end of each flight. The circle with its arrows has the same dimensions and the same appearance as the circle with the single arrow described in "e." The outside diameter of the circle is 0.25 inch and the length of each arrow, extending to the center of this circle, is 0.19 inch The letters at the end of the arrows should be as large as possible without interfering with the perception of the diagram.

TECHNICAL ASPECTS OF THE PRODUCTION OF FILM

A step-by-step account will be given of the production of psychological test films from the time the final specifications are drawn up to the completing of the finished product in the form of 16 millimeter release prints.

A motion picture film, if more than one copy is desired, is subject to the photographic principle that a positive can only be made from a negative and a negative can only be made from a positive. The film that is projected—the only film that can be projected—is positive film. A motion picture negative is never projected. Because it is easily scratched, it is handled only by technicians trained for that purpose, who wear special gloves and observe a number of other elaborate precautions. Negative film is, of course, the first product of motion picture photography. Its sole use is to be run through a printing machine which makes positive prints.

At the time that the negative is exposed in the camera, i. e., when the specified shots are taken, some footage is run off at the end of each roll of film. Its chief purpose is to make a Cinex (Sinex) Test which determines the intensity of light which will be necessary in printing to insure a uniform density in the positive prints. In a professional studio, at the end of each day's work, all the exposed negative is taken to the laboratory. Here the negative is developed and positive prints are made from it. Laboratory work is usually done at night so that the positive prints can be viewed the day after the film was shot. These positive prints, made from the negative exposed the previous day, are known as "dailies" or "rushes." The dailies are viewed by the editor or cutter and all others immediately concerned with the production. While the film is being projected, criticism is made of its quality and effectiveness. At this time, selection of "takes," or alternative shots, are made, since usually more than one take has been made of each specified shot. The cutter makes notes of decisions and selections.

After the dailies have been viewed and judged, the cutter takes the positive film to the cutting room. Here, he breaks the film down into the individual takes or scenes. At this time too, the "key numbers" of the film (numbers printed on the margin of the film at each successive foot of its entire length) are recorded and filed. This is done so that in the event the cutter has to call for a reprint of any part of the original negative he can do so by ordering the reprint from number to number. This eliminates the necessity of reprinting a whole sequence when only a portion of it is needed.

Using the list of specifications and the script as guides to the continuity, as well as the notes made during the running of the dailies, the cutter arranges the takes in their proper sequence at

the proper intervals. The cutter's main tool is the Moviola. This is in essence a miniature projection machine. The great advantage of the Moviola over the regular projection machine is that the former can be slowed up or reversed and can be stopped anywhere the operator desires, precisely to the frame. Thus the cutter can run the film on a take, back and forth, until he finds the exact frame in which he wants to make his cut. The amount of a given shot that the cutter uses depends partly on his judgment of the continuity of action in going from one scene to another, partly on the script, and principally on the amount of footage required for the amount of narration written for that scene. Standard 35 mm. film is photographed and projected at 24 frames (one and a half feet) per second, or 90 feet per minute. Narration is usually spoken at about 145 words per minute, and if pauses are required, as in the case of test directions, their duration is specified; hence the duration of the action and of the corresponding voice can be matched to one another in terms of film footage. When the scenes have all been arranged and the cuts made as desired, the work print is spliced together.

Splicing of film, although quite simple with the proper tools, is nevertheless very important. The splice must be smooth and durable so that it can go through the carriages of either the projection machine or the moviola and draw the rest of the film after it. The splicing machine is a device that uses pressure and heat. The emulsion is removed from the end of the strip to be spliced, a special film cement is applied with a thin brush, and the two ends are brought together so that they overlap the width of the frame mark, being automatically trimmed at the same time. The ends rest between two irons that are heated electrically. The splice is allowed to remain in this position for a few seconds and then the irons are lifted off. The film should then be securely and evenly spliced. Usually at this stage the film test is tried out with a group of subjects and revisions are made.

Commercial entertainment pictures have the sound recorded at the same time that the picture is taken and they are automatically synchronized. This is, of course, necessary where dialogue accompanies action and the screen must show the actor speaking. The unseen narrator is, however, coming into increasing use, especially in instructional films, and in this case the action and the voice need not be synchronized literally but instead need only be coordinated logically. The action explains or exemplifies what the voice is talking about. Narration of this sort is recorded on a separate film after the picture proper is cut and spliced. The picture and the sound track are later combined into a single film. This was the procedure employed for the psychological test films, in which the voice is simply that of a disembodied test administrator

explaining the task portrayed and giving directions as to how to proceed.

After the picture is spliced, the narration as written in the revised script is recorded at an average of three words to two feet of film. This is done in a sound-proof room. The narrator first rehearses the prepared script for the designer of the test until the reading is given with the desired emphasis. When the narrator can read the script to the satisfaction of both the test designer and the sound technician, the recording is made. Complete sentences or paragraphs are recorded as takes, the narrator pausing and repeating if mistakes are made. The sound track will later be cut and edited just as the picture was. The track is recorded on film which is processed through the laboratory in the same way as picture film.

The sound recording is judged, in the form of dailies, and selections are made by the cutter and others. The cutter takes the film bearing the track and breaks it down so that the sentences and phrases correspond roughly to the action called for in the script. With the aid of a synchronizer and the moviola, he cuts and matches the film carrying the sound to the film carrying the action. The cut sound track is then spliced. Now the cutter has the picture and the sound track, each on separate reels.

By the use of projectors designed to run the separate picture and track films in synchronization, the cutter views this matched work print as it will appear on the screen. If there are corrections to be made, the cutter takes the film back to the cutting room and makes them. The matched workprint then meets the requirements called for in the script.

If the test film requires any special effects photography such as dissolves, fades, superposures, or other "optical work," the cutter "cues" the picture by writing on the film with a grease pencil, and orders the necessary special photographic treatment of those particular sequences of the film. This involves the making of new positive, new negative, and positive again for those sequences. The sections of new negative are made on an optical printer which is capable of producing a variety of special effects on the original action filmed. After inspection of the new positive film, the cutter substitutes the special effects for the cued parts of the film. At this stage of the project, the work print is in its tentative final form. It is now viewed by the psychologist and the cutter. If approved by the former, the test film is ready for negative cutting.

The original negative, including the special effects negative, has scarcely been handled during all this time. It has been labeled and filed in vaults, reel by reel. It is now taken out and, with the spliced work print as a guide, is lined up to correspond to the positive. With the aid of a multiple-headed synchronizer for

matching one film to another, the negative is cut so that it has a cut wherever there is a splice in the positive. This will produce a negative that corresponds exactly to the work print, with leaders, titles, special effects, and picture included. The cut negative is now very carefully spliced. At the same time, the corresponding sound track is also matched with the sound track of the work print and is also spliced.

From this spliced dual negative (both picture and track) the laboratory prints a "master composite positive." For the first time, both picture and track are combined on one continuous strip of film. From this master composite is made a 16 mm. sound negative, by reduction printing, and from this negative are made the 16 mm. prints which constitute the copies of the test. The original 35 mm. negative and the master positive are preserved in film storage vaults.

CHAPTER FOUR

The Presentation of Motion Picture Tests and Other Films Requiring Activity by the Group*

Group testing with motion pictures is practicable only if two objections, which will occur to anyone who thinks about the matter, can be answered, namely, will some of the individuals being tested have the handicap of a poorer view of the screen than others? and, how can they see to write their answers when the room is dark? Satisfactory answers to these questions can be given, as this chapter will try to show. A film can be seen from a large number of viewing positions which are demonstrably equivalent. And, perhaps surprisingly to some, it is not necessary to darken a movie room but only to dim it. The evidence on which these answers are based will be given. The implications of this evidence extend to the use of film for *all* educational purposes as well as for testing.

In the course of administering tests on film to some thousands of aircrew candidates and students, the Psychological Test Film Unit made a continuous study of the viewing circumstances under which these films could be shown and of the effect of these different circumstances on the performance of the students as measured by their test scores. This research was necessary in order to determine, first, whether these circumstances, such as seating position in the testing room, made any difference in the success with which the visual problems of the test were solved and, second, what circumstances were advantageous for the presentation of motion picture tests and examinations. A large mass of evidence was collected on the effect of viewing position and illumination on test performance, and the accumulation of experience led to a set of opinions on what a motion picture viewing room should be.

THE PRESENTATION OF FILMS FOR PURPOSES OF INSTRUCTION OR TESTING

It is obvious that a test film cannot be given without some method of illumination in the room by which the student can

*This chapter was prepared by Ben C. Finney and the editor.

write. He must at least be able to make marks on an answer sheet in such a way that they can later be scored. In the case of a test like the examination in practical aerial navigation to be described later, he must be able to make computations and study maps while still being able to see the screen clearly. It is probably less obvious but equally true that an instructional film should be presented under the same kind of special illumination of the room. The student ought to be able to take notes, sketch diagrams, or study pictures in a book while the film is being run. Learning is enhanced by activity; the student who sits back and listens passively in a classroom is ordinarily not doing much learning.

The technique of showing films in a semi-illuminated room has been used many times and with many subjects who have taken motion picture tests. It could be used equally well for teaching with films. But it is not now used for that purpose in most schools and colleges. It appears unorthodox and is completely at variance with the traditional situation for viewing motion pictures—a dark room, suggesting relaxation and escape from reality, which encourages the onlooker to lose himself in the events being portrayed or narrated by the film but does not permit him to write or take notes, to see the instructor in the room or to interject a question.

The attitude of the ordiary moviegoer is quite different from that of the student. The former is seeking entertainment and has no special intention of remembering what he sees and hears. The student is actively learning. If his intention is to remember the material being taught, he uses a notebook and a pencil. The entertainment screen situation and the educational screen situation produce their own unique attitudes in the onlookers. One of the causes of this difference in attitude is the illumination of the room. Darkness is notably a condition which promotes a tendency to rest, and a non-critical attitude. It is appropriate for the psychoanalyst's consulting room, for daydreaming, and for storytelling. The darkness of the movie theater is right and proper for appreciating the make-believe world which opens up on the motion picture screen. But it is not right for a classroom with students and a teacher who need to see one another. Since a semi-illuminated room is seldom employed at present in showing educational films, the attitude of those who see them is too frequently the one which they adopt as moviegoers—the attitude of a resting spectator.

During the war many thousands of army students sat in dark rooms and saw many hours of training films. They were frequently shown during the latter part of a physically active day and the trainees were tired. An attitude of passive relaxation was only to be expected. The training films themselves were in many cases conducive to effective learning, but the conditions under which they were shown were not. It is our hypothesis that the "spectator atti-

tude" in viewing such films can be combatted by setting up a viewing situation appropriate for active learning—a situation similar to that existing in a good classroom, and to that set up for motion picture tests. It is an even more obvious hypothesis that only in such a viewing situation can the student take notes on what he is being taught or write answers to questions which are put him.

Fundamentally the requirements of a room in which motion picture testing or teaching is to occur are: *proper illumination* in which to take notes and to see the instructor, and *proper seating and projector arrangements* which will permit perception of the screen without effort.

Desirable Features of a Classroom in Which Motion Pictures Are To Be Used for Testing or Teaching

Illumination. The viewing situation for the presentation of films for educational purposes, including that of testing, is very much easier to establish than is ordinarily assumed. The semi-illuminated classroom is already used for the projecting of lantern slides, particularly the lecture-table types where the image is thrown on a screen above the instructor's head. It is just as appropriate for educational films. It would permit a much more flexible use of films than is at present customary—a use appropriate to instruction rather than to storytelling or "narration," and one which does not relegate the classroom instructor to the status of a nonentity while the film is being shown.

The system of illumination used and recommended by the Psychological Test Film Unit for giving motion picture tests was a set of low-wattage ceiling lights controlled by a separate switch from that which provided the full normal illumination of the classroom or testing room. These low-intensity light bulbs, with or without diffusing fixtures, should be shielded in such a way that they do not throw light directly on the screen. Shields of black cardboard are satisfactory for temporary purposes. The ceiling should be high enough so that the lights do not fall within the normal field of vision of the students. Windows, of course, should be covered with shades. If ventilation is dependent on the opening of windows, a simple type of louver covering the upper half of the window area may be used as a light-trap which will at the same time permit the upper window to be kept open. This was the system used at Santa Ana Army Air Base for motion picture testing and also for the classroom teaching of aircraft recognition which required a high proportion of class time to be spent on visual instruction using lantern slides. It was a permanent setup for illuminating the rooms which required it. Temporary arrangements are, of course, possible merely by substituting low-wattage bulbs in the permanent lighting fixtures of a room. The above method proved to

furnish adequate light for marking answer sheets and taking tests.[1]

Other systems of semi-illuminating projection rooms are possible. The projection rooms of professional Hollywood studios, where studying of a certain type is required, are sometimes fitted out with desks having indirect illumination of their writing surfaces. A certain amount of glare is almost unavoidable, however, even from the shielded lamp fixtures used. A better system is focussed pencils of light from ceiling fixtures which shine on the individual desk surfaces. This method would be expensive for a large classroom. The simpler method described above is believed best.

The question of what *level* of illumination on the writing surfaces of the individual students is best, is a matter for experiment. Evidence will be presented to show that it may vary widely without affecting the performance of the student. Obviously, there are two contradictory aims in the illuminating of classroom projection rooms. The more light in the room, the better the students can write, see the instructor, and keep alert. On the other hand the more light the less is the degree of "contrast" of the screen image and the poorer therefore its photographic quality. Illumination falling on the screen will tend to "wash out" the blacks of the image. This effect, however, seems to be less than might be theoretically supposed, and it is possible that the mechanism of brightness constancy is at work to preserve the appearance of screen images at different illumination levels. At any rate, a compromise may be reached practically. The experience of the Psychological Test Film Unit indicates that the student may be

---

[1]A memorandum recommending the above arrangement may be quoted here as having historical interest:

"1 April 1944. SANTA ANA ARMY AIR BASE.

"SUBJECT: Illumination of Aircraft Recognition Classrooms.

"TO: Department Chairman, Aircraft Recognition, Preflight School.

"1. The present illumination of classrooms used for Aircraft Recognition has often proved a hindrance to effective instruction. The Psychological Test Film Unit, in connection with its own program, has developed an ideal illumination situation which it recommends to the Aircraft Recognition Departments for their use in classrooms.

"2. The recommendations are made with the knowledge that classrooms are often used by other departments requiring fuller illumination than is desirable for Aircraft Recognition needs.

"3. These are the recommendations:

a. Complete "blackout" for windows which are now only "grayed-out" or partially obscured; continued use of ventilator louvers as in the past.

b. Substitution of 15-watt globes in *one circuit* of lights for the present ones of high wattage; continued use of globes of high light intensity in other circuit.

c. Use one circuit with the 15-watt globes for Aircraft Recognition instruction; use both circuits for other classes requiring fuller illumination.

"4. The above plan for Aircraft Recognition classroom illumination should result in uniform lighting in recognition classrooms, adequate light to take notes and record written responses to tests and drills without detracting from sharp projection of images on the screen, and elimination of complaints of 'too much light,' 'not enough light to write by,' etc."

given from two tenths to eight tenths of a foot-candle of illumination to write by without seriously affecting the visibility of the screen.

Setup of the Projector and Screen. The visibility of the screen image is affected by the length of throw of the projector, the type of projector, the intensity of the projector lamp, and the type of the screen. No experimental evidence has been obtained on the effect of these variables. However, application of certain optical principles together with information from other studies makes possible some generalization concerning the setup of the projector in the room. Since the brightness of the screen image varies inversely with the square of the distance of the projector from the screen, the closer the projector to the screen the brighter and smaller the image. Thus, if the brightness of a projected image does not seem adequate for good visibility, it can be increased by moving the projector nearer the screen. However, the screen image becomes smaller. The effect of this decrease in the size of the screen image can only be considered in relation to the size of the projection room and the seating distances from the screen of the students in it. The question of how much variation in seating position is possible without handicap to some students will be discussed in the next section. The size of the screen image, in general, should be as large as the screen will permit, taking account of the desired brightness of the image.

Since, in a partially-illuminated room, the stronger the light thrown on the screen the greater will be the contrast of the image, it should be as bright as possible for a given size of the image. Therefore the projector should be fitted with a lamp of maximum intensity.

Of the various types of screen currently available the white matte screen has proved to be the most suitable. Beaded screens are unsatisfactory for the reason that brightness of the screen image is noticeably less when the screen is viewed from the side of the room than when seen from the center.

Seating Arrangement. Since all the individuals taking a motion picture test must have an equivalent view, although not necessarily an identical one, of the pictures shown if their test scores are to be fair measures of ability, the equivalence of different seats in a testing room must be demonstrated. Viewing positions too near or far from the screen or too far toward the side must be eliminated, or else the scores at these positions must be statistically corrected. The effect of seating position on test scores was therefore studied with all the motion picture tests constructed by the Psychological Test Film Unit. Whenever a test was administered, a plot or map of the seats in relation to the screen was made out and a mean score for each seat or block of seats was computed. In addition,

the students in the presumably less favorable positions were often asked individually whether their view of the screen was satisfactory.

The results of a long series of such studies, presented in the next section of this chapter, showed that students can sit at distances and angles of view which appear to be highly unfavorable to efficient perception of the screen *without any significant lowering in the level of their performance.* It might be supposed that test performance in unfavorable viewing positions was compensated for by some kind of extra effort of the subjects; there was, however, little or no evidence of such effort from the replies to individual questioning. A more probable explanation is that the different visual stimuli affecting the eye at different viewing positions are all equivalent for performance. Within wide limits, it appears that different seats in a projection room provide equal opportunity for perception.

Recommendations for the seating of students in the educational use of films, on the basis of this evidence, may therefore be very loose. Visual education authorities have recommended, without experimental facts to support them, the following rules: The nearest row of seats to a screen should not be closer than two picture-widths, and the farthest should not be more distant than six picture-widths. The angular retinal size of the picture, therefore, should lie between 29° and 10°. The boundaries of seats at the side of the room should be set at 30° on either side of the perpendicular to the screen, or the axis of the projector's beam, which would provide a wedge-shaped block of seats, 60° wide. These rules, in relation to our results, are more than safe. In the light of the evidence to be reported, they could be somewhat exceeded without significant unfairness to the students.

Ventilation and other Factors. Only the same precautions ordinarily observed in good classroom teaching need be followed in using motion pictures for testing purposes, and also for instructing. Since windows must be darkened by shades, either forced draft ventilation or the system of window-louvers mentioned is necessary.

Proctoring of Tests and Examinations. It might be assumed that, in semi-darkness, the subjects of a motion picture test will be tempted to copy their neighbors' answers. Although all subjects are, it is true, reacting to the same questions at the same time, and are seated side by side, the pace of the test is such that there is insufficient time to locate and copy a neighbor's answer before the next question appears on the screen. It has been the experience of this organization that the problem of copying answers is almost eliminated in motion picture testing and that proctoring is minimized. In certain kinds of motion pictures tests in-

volving the "when to react" type of discrimination it is necessary to enforce the rule that no individual start to mark his answer until all do at once at the beginning of the answer period. Otherwise social cues might influence the choice of an individual's response among the sequence of alternatives.

EXPERIMENTAL EVIDENCE ON THE EFFECT OF SEATING AND ILLUMINATION ON TEST SCORES

Some 30 or more experimental studies were conducted to determine whether the viewing position of the student in the motion picture testing room, and the level of illumination of the room by which to record answers, affected his performance on the test. Of these studies, 20 are available for summarization in this report. They were performed with ten of the motion picture tests to be described in the next two chapters. In the case of a few tests completed and tried out toward the end of the war period, the data were not computed in time for inclusion here. The evidence is sufficient, however, for general conclusions.

The various aptitude and proficiency tests were given to aviation students in three types of rooms: (a) rooms of classroom size in the neighborhood of 35 by 25 feet, and seating 24–40 students; (b) large rooms designed for group testing, filled with desks or shallow cubicles, approximately 60 by 60 feet in size and seating 180–200 students; and motion picture theaters with a large screen (and a corresponding long throw of the projector) seating students at distance up to 120 feet from the screen and in numbers up to several hundred at a time. All three of these situations proved to be practicable for motion picture testing. Use of a theater required that lap-boards be used for writing, which was inconvenient; the theater lights, however, could be adjusted by dimming to any level of illumination desired. Students had to be seated in alternate rather than adjacent seats since, in a theater, the latter are so close together. It was concluded that testing in rooms smaller than a theater was somewhat preferable, although not decisively so.

The studies of viewing position and illumination made in these situations were of two types: (a) studies in which all individuals being tested had a "favorable" view of the screen (no seat nearer than two picture-widths nor farther than six picture-widths from the screen, and no seat at a greater viewing angle than 30° from the center line), and (b) studies in which some individuals had "unfavorable" viewing positions at the extremes of distance and at extreme angle of view. The latter studies were performed in the effort to discover at what distances, if any, and at what angles of view, efficiency of perception would begin to decrease in a motion picture room. In the former studies the illumination of

the room was at an estimated optimum; in the latter studies the illumination was sometimes varied, with different groups, from a wholly dark room to a completely illuminated room.

The procedure of the experiments was to make a seating plot of the testing room, each seat being indicated by a letter and number combination. Viewing positions were divided into convenient blocks according to distance from the screen and also according to angular deviation of the line of sight from the straight-front line. Each student entered the seat number on his answer sheet. When the test had been administered to a sufficient number of groups, the papers were sorted first for distance and then for angle, means and sigmas being computed for the various blocks of seats. In the majority of studies differences and their critical ratios were calculated, but in seven of the studies use was made of the analysis of variance technique.

The Effect of Distance and Angle of View

A detailed presentation of the 20 seating studies completed is impossible within the limits of this chapter. They are given in the *Annual Report* of the Psychological Test Film Unit for the fiscal year 1945. Instead, one seating study will be described as an example, and the results of the others will be tabulated. With almost no exception they lead to the same general conclusion, namely that most of the tests can be viewed from positions up to highly "unfavorable" extremes without any significant falling off in test performance, but that a few tests cannot. In the case of three tests, involving discriminations analogous to those of visual acuity, performance falls off with increasing distance, but not with increasing angle. For these tests, and these alone, the effect of distance on test scores does not wholly disappear even when the distance from the screen of the nearest and farthest seats is within the limits which have been termed "favorable."

The Effect of Viewing Position on Scores of the Estimation of Relative Velocities Test, CP205B-III

The experiment to be described is one in which the viewing positions were extended to the most "unfavorable" extreme limits of angle and of nearness to the screen. Some seats in front and at the side were as much as 80° from the straight-front axis. The room was one used for group testing and, by design, the seating was not adapted to motion picture presentation. The front row of seats was less than four feet from the screen. The size of the room was 49 by 57 feet, with seats extending to the walls on all sides and nearly all the way to the front wall. A less favorable situation for viewing motion pictures could scarcely be set up. The

test was given at Nashville Army Air Center to 1,104 unclassified students, in groups of about 200, in November 1943.

Effect of Angle of View. The viewing positions were divided into seven sectors, as shown in figure 4.1, representing angles of the line of view within 15° of the center line, and from 15° to 30°,

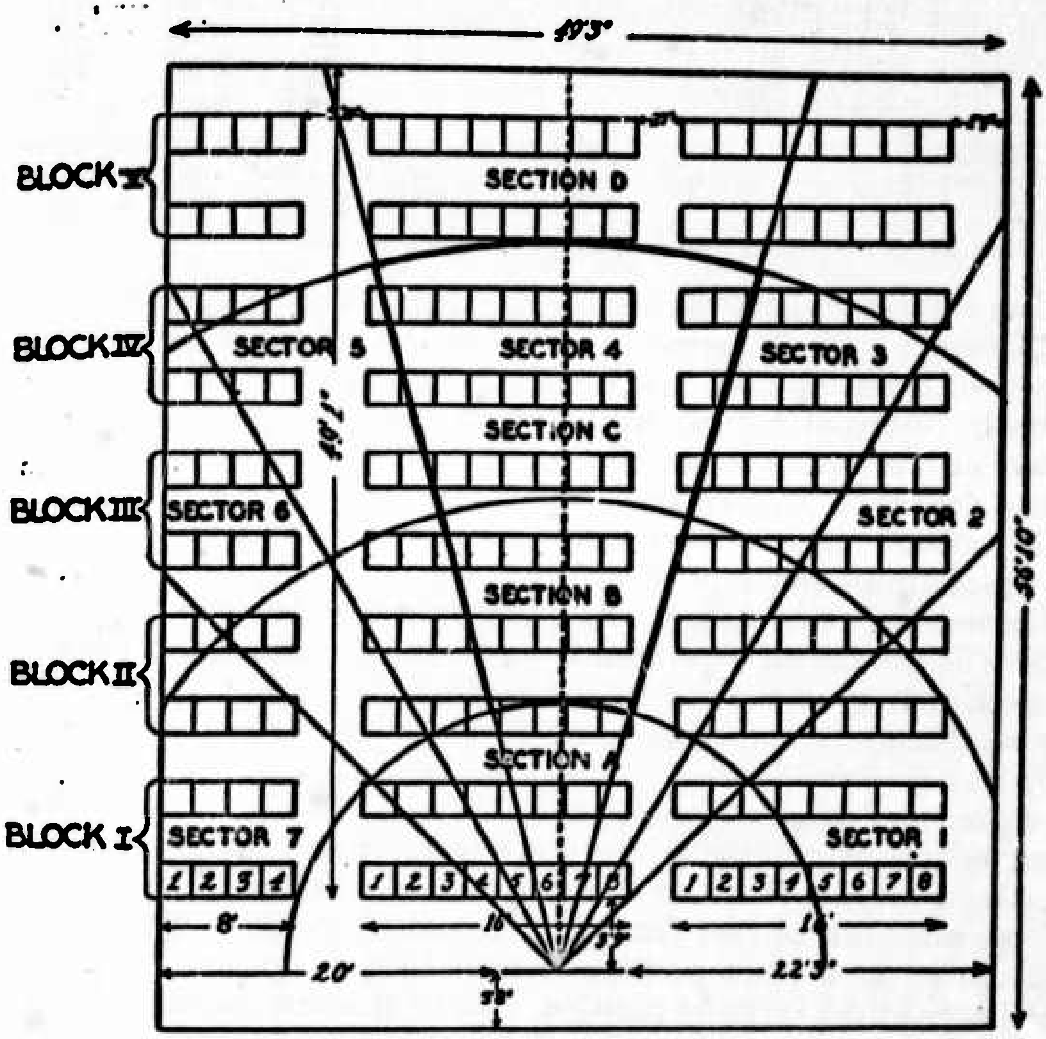


FIGURE 4.1.—Plot of Group Testing Room, Nashville Army Air Center.

from 30° to 45°, and from 45° to 80° away from the center line on either side. The mean score was computed for students in each sector and critical ratios were obtained for comparisons between sectors. In addition, comparisons were made for the group of students seated in the *middle range of distance* within the central sector, i. e. in the most favorable positions. This group is labeled the "optimal center" and its limits are indicated in heavier lines on the seating plot. The seats were from 12 to 40 feet from the screen and at angles within 15° of the center line. The results are given in table 4.1, and the critical ratios of the differences are listed on following page.

Table 4.1.—*Test scores in relation to angle of view*

Sector		N	M	SD
1	(45°L to 80°L)[1]	101	27.91	5.03
2	(30°L to 45°L)	101	29.98	4.64
3	(15°L to 30°L)	194	30.44	5.42
4	(15°L to 14°R)[1]	347	29.74	5.19
5	(15°R to 30°R)	167	30.12	4.97
6	(30°R to 45°R)	94	30.49	5.55
7	(45°R to 80°R)	64	31.00	5.94
	Optimal Center	235	30.20	5.37

Sector	C.R.	Sector	C.R.
1 *vs.* Optimal center	3.70	1 *vs.* 4	3.20
7 *vs.* Optimal center	1.19	1+7 *vs.* 6+2	1.27
1+7 *vs.* 4	0.60	1+7 *vs.* 3+5	1.49
2+6 *vs.* 4	0.90	5+6+7 *vs.* 1+2+3	1.47

[1] L—Left side of room facing screen; R—Right side facing screen.

The differences between sectors are slight. The only significant difference is between sector 1 and the central positions. As can be observed from the plot, many individuals in sector 1 had an extremely distorted retinal image of the screen. The equivalent group of individuals in sector 7, however, did *not* show any significant decline in performance. The N for sector 7 is smaller, and it is probable that the very worst seats at an 80° angle were not occupied in all the repetitions of the experiment. There is no difference between central and peripheral seats in general, and scores do not fall off significantly until the viewing angle becomes considerably greater than 45°. It may also be noted that there is no significant difference between the two sides of the room (1, 2, and 3 *vs.* 5, 6, and 7).

Effect of Distance From Front of Room. The extreme in the near relationship to the screen was well represented in this study, positions varying from a distance of 3 feet 9 inches from the plane of the screen to 50 feet from it. The retinal image of the screen picture at the nearest position is more than 10 times the size of the image at the farthest position. Blocks of seats, combining two rows, were numbered from I to V, as shown in the plot. Means, sigmas, and critical ratios were obtained, and are given in table 4.2.

Table 4.2.—*Test scores in relation to distance*

Block	N	M	SD
I (3' 9" to 12')	211	29.21	5.37
II (12' to 21')	234	30.21	4.96
III (21' to 33')	236	30.30	5.31
IV (33' to 40')	201	29.81	5.50
V (40' to 50')	191	29.68	5.08

Block	C.R.	Block	C.R.
I *vs.* II	2.03	I *vs.* V	0.90
II *vs.* III	0.19	II *vs.* V	1.08
III *vs.* IV	0.52	III *vs.* V	1.23
IV *vs.* V	0.29		

The scores in the first two rows are slightly but not significantly lower than the scores farther back in the room. These front seats are subjectively very unfavorable viewing positions, including as they do both seats too close to the screen and seats at too great an angle.

Effect of Distance From Center of Screen. A better comparison is between different distances from the screen rather than distances from the front of the room. Arcs were drawn on the plot at radii of 16, 28, and 44 feet, dividing the room into four sections lettered from A to D, as illustrated in figure 4.1. Means and critical ratios were computed, as before. The results are given in table 4.3. No differences are significant. Although nearly all the

TABLE 4.3.—*Effect of distance from center of screen*

Section	N	M	SD
A (3' 9" to 16')	195	29.83	5.17
B (16' to 28')	318	30.33	4.94
C (28' to 44')	364	29.98	6.78
D (44' to 56')	200	29.67	5.18
Section	**C.R.**	**Section**	**C.R.**
A vs. B	1.08	A vs. C	0.29
B vs. C	0.77	A vs D	0.31
C vs. D	0.61	B vs. D	1.44

seats in section A are nearer the screen than two picture widths, they do not produce a real handicap in taking the test. For this particular test (ability to estimate relative velocity), and for most of the other tests, the size of the retinal image of the picture does not seem to affect performance. Other studies have explored the effect of distance to *far* distances, which are not represented in this experiment. Up to 100 feet, no effect is evident for the majority of tests. Other studies have also shown insignificant variance due to angle of view, at less unfavorable extremes than were here tried out. One study explored the possibility of *interaction* between the isolated effect of distance and the effect of angle. No interaction appeared.

Effect of Right-Left Relationship. In order to determine further whether the right and left sides of the room were equivalent for test performance, one final comparison was made. The three extreme seats on each side of the room, from front to back, were combined. The results are given in table 4.4. The difference is not significant.

TABLE 4.4.—*Effect of right or left side of room*

	N	M	SD	C.R.
Right side	159	30.24	5.52	0.98
Left side	166	29.70	4.87	....

The results of the above experiment represent the usual outcome of seating studies made with tests which *do not require fine visual discriminations of an absolute type,* such as visual acuity.

Summary of Seating Studies on Distance and Angle

A tabular summary of 20 seating studies is presented in table 4.5. The effect of distance and the effect of angle are listed as "significant" or "not significant," depending on whether the critical ratios (or F-ratios) met the criterion of statistical significance

(1% level). "Extreme" distance in the table means an angular retinal size of the picture (visual angle) smaller than 10°. This angular size takes into account both the distance of the observer from the screen and the size of the screen image. "Non-extreme" distance means a retinal image larger than 10°. "Extreme" angle of view means a viewing position more than 45° from the center line. Since some studies were not set up to explore the extreme seating angles and distances, the notation "No evidence" is frequent in the table. The effects listed as "significant" are printed in italics. In all these twenty studies the illumination of the room was nearly constant and was close to "optimal."

It may be noted first from the last two columns of the table that, with the exception of the experiment described above, there is no significant effect on test scores from the angle at which the screen picture is viewed. This fact holds for all tests. Only two studies included "extreme" viewing positions at angles greater than 45°. The first has already been described. The second yielded no significant effect, with angles up to 62°. The results indicate that perceptual efficiency in viewing pictures is not affected by the viewing angle within a much wider range than is ordinarily supposed.

The columns headed "Effect of Distance" present a different state of affairs. Three tests show significant decreases in test performance with increasing distance from the screen, the Flexibility of Attention Test, the Integration of Attention Test, and the Minimal Movement Test. (The one significant effect under the Plane Formation Test will be considered later.) In the case of all three of these tests, the difference in test performance shows up at nonextreme distances as well as at extreme distances. The fact is that in these tests, performance is best when close to the screen and becomes slightly but consistently poorer as the screen image becomes more distant and the retinal image becomes smaller.

These tests are those, out of all tests constructed, which require judgments in which the resolving power of the visual mechanism seems to be important. In the case of the Flexibility of Attention and Integration of Attention Tests the observer must watch a number of dial hands or indicators which may move into a region marked in black (i. e., may go "wrong") or may simply move very close to that region. The discrimination between the coincidence or non-coincidence of the indicator sometimes requires a high degree of visual acuity. If acuity is dependent on optical and retinal factors, it is reasonable to suppose that the discriminations would become more difficult as the angular retinal size of the indicator becomes smaller. In the case of the Minimal Movement Test the observer must distinguish between barely perceptible movement of a black spot and no movement. The absolute thresh-

TABLE 4.5.—*The effect of differences in viewing position on mean test scores*

No. of subjects	Type of projection room	Min. distance from screen and visual angle	Max. distance from screen and visual angle	Max. angle at which viewed	*Effect of distance from screen on mean test scores* Nonextreme (over 10°)	Extreme (under 19°)	*Effect of angle of view on mean test scores* Nonextreme (under 45°)	Extreme (over 45°)
				Estimation of Velocity CP205B-I				
450	Theater........	44' (14°)	71' (8°)	20°	No evidence........	Not significant.....	Not significant.....	No evidence
407	Classroom......	16' (23°)	32' (12°)	20°	Not significant.....	No evidence........	....do.............	Do.
				Identification of Velocities CP205B-II				
450	Theater........	44' (14°)	71' (8°)	20°	No evidence........	Not significant.....	Not significant.....	No evidence
407	Classroom......	16' (23°)	32' (12°)	20°	Not significant.....	No evidence........	....do.............	Do.
				Estimation of Relative Velocities CP205B-III				
450	Theater........	44' (14°)	71' (8°)	20°	No evidence........	Not significant.....	Not significant.....	No evidence
407	Classroom......	16' (23°)	32' (12°)	20°	Not significant.....	No evidence........	....do.............	Do.
1,194	Test room......	4' (82°)	56' (7°)	80°	....do.............	Not significant.....	....do.............	Significant (in part
1,160	Classroom......	13' (21°)	25' (12°)	16°	....do.............	No evidence........	....do.............	No evidence
1,430	Test room......	10' (38°)	75' (4°)	62°	....do.............	Not significant.....	....do.............	Not significant
				Plane Formation CP805B				
450	Theater........	44' (14°)	71' (8°)	20°	No evidence........	Not significant.....	Not significant.....	No evidence
407	Classroom......	16' (23°)	32' (12°)	20°	*Significant*.........	No evidence........	....do.............	Do.
				Flexibility of Attention CP411B				
1,754	Test room......	23' (26°)	108' (6°)	45°	No evidence........	*Significant*.........	Not significant.....	No evidence
1,000	....do.........	23' (26°)	108' (6°)	45°	....do.............	....do.............	....do.............	Do.
1,378	Classroom......	11' (24°)	29' (11°)	35°	*Significant*...?.....	No evidence........	....do.............	Do.
				Integration of Attention CP411D (and CP415A)				
1,000	Theater........	23' (26°)	108' (6°)	45°	No evidence........	*Significant*.........	Not significant.....	No evidence
1,378	Classroom......	11' (24°)	29' (8°)	35°	*Significant*.........	No evidence........	....do.............	Do.
				Minimal Movement CP213B				
168	Classroom......	10' (29°)	30' (9°)	25°	*Significant*.........	No evidence........	Not significant.....	No evidence
				Drift Direction CP221A				
168	Classroom......	10' (29°)	30' (9°)	25°	Not significant.....	No evidence........	Not significant.....	No evidence
				Successive Perception CP509C-II				
381	Classroom......	18' (30°)	87' (11°)	18°	Not significant.....	No evidence........	Not significant.....	No evidence
				Aircraft Recognition Proficiency Test				
250	Classroom......	10' (29°)	30' (9°)	45°	Not significant.....	No evidence........	Not significant.....	No evidence

old for movement is analogous to acuity in being dependent on the angular or retinal extent of the displacement, as distinguished from the perceptual size of the total presentation. It is therefore dependent, at least in part, on the viewing distance. This interpretation is borne out by the fact that the Drift Direction Test, employing the same spot, crosshairs, and circle, but requiring a judgment of the inclination or drift of a movement rather than the existence of a movement, does *not* show any decrease in performance due to viewing distance.

These three tests account for all the significant effects of viewing distance found in all the studies. There is one exception. One of the two studies reported for the Plane Formation Test yielded a significant decrease in scores with increasing viewing distance. The effect was obtained with moderate distances. Doubt is cast on this result by the fact that another study, at extreme distances, showed no effect at all. The result is rendered even more dubious by the circumstance that the study in question was incomplete, having been carried out by another research unit, and that the standard deviation for the scores had to be estimated by this organization. The critical ratio of the difference, when computed, turned out to be significant, but it is highly questionable. It appears unlikely that performance on the Plane Formation Test is affected by seating-distance from the screen.

Subject to the limitations discussed, it may be said that perceptual efficiency in viewing motion pictures is not affected by the distance of the observers within the range of the ordinary classroom or small theater. If a film requires a visual task which is likely to make demands on visual acuity, special seating arrangements will have to be made. But for most types of visual tasks, all the observers in a movie room have substantially an equivalent perception of the screen as regards performance.

The Effect of Illumination of the Room

A number of informal tryouts and two formal experiments were carried out on the effect of the level of illumination on performance with motion picture tests. It was discovered that a level of illumination could be obtained, either by using low-wattage light bulbs or by adjusting the brightness of ceiling lights with a rheostat, which would be subjectively "satisfactory" for the taking of motion picture tests. At this illumination, students stated that they could write answers comfortably, and the screen image was judged by the psychologist to have little loss of contrast. It was in the neighborhood of two-tenths of a foot-candle at the writing surfaces used by the students. A considerable variation in the illumination above or below this point seemed, however, to be

possible. In order to verify this conclusion, two experiments were carried out.

The first of these compared the mean scores obtained for two groups of students on the Estimation of Relative Velocities Test, using a group testing room, under ordinary dim lighting (N=1480) and under "blackout" conditions (N=242). The first condition was approximately "optimal"; the second involved an almost completely dark room in which the answer sheets were barely visible. The mean scores were 28.46 and 28.16 respectively and the critical ratio was .80. The extreme of low illumination did not affect test performance.

The second of these studies analyzed the variance due to illumination in two different conditions of the Flexibility of Attention Test, using a theater. The first condition was dim lighting, at a level of approximately one-tenth of a foot candle (N=845). The second was full illumination, with the theater lights at their highest brightness, at a level of approximately 1.4 foot candles (N=909). The differential was 14 to 1. In the latter condition the screen image showed a noticeable loss in contrast, and the illumination was definitely not "optimal." The F-ratio between illuminations was, however, not significant at the 1 percent level, although it was significant at the 5 percent level. The extreme of high illumination did not, to a formally significant extent, affect test performance.

These studies demonstrate that, for motion pictures requiring "audience participation," a wide variation in level of illumination is permissible without any overt effect on perceptual efficiency. An estimated optimal illumination may therefore be set up without fear of being inexact in attaining it. The simple expedient of low-wattage ceiling lights is fairly sure to be satisfactory.

CHAPTER FIVE

Aptitude Tests*

BACKGROUND OF TEST RESEARCH

A number of assumptions governing the research on aptitude tests with films were listed in Chapter 1, and they may be repeated here, after intervening chapters on the nature of the motion picture medium, the making of motion pictures, and the showing of them for educational and psychological purposes. It was assumed that:

1. Visual and spatial performances being important for fliers, these functions should be concentrated on in devising motion picture tests.

2. In devising a test, the procedure should involve not only job analysis of the performance but also psychological analysis of the underlying function in terms of the theory of perception.

3. The methods of psychophysics in studying simple and complex *discriminations* by means of repeated *judgments* are useful in setting up the test, once the kind of discriminations required have been decided upon. A motion picture test can be built on the pattern of a psychophysical experiment, which involves trials rather than items. In such an experiment, unlike a printed test, the same presentation can be repeated a number of times in a random series.

4. The functions tested should be those to which the medium is uniquely adapted.

5. These functions will probably be found to center around the human capacity for locomotion through space, and on that account will require a theory of continuous behavior rather than a simple theory of stimulus-response behavior.

Employing these more or less implicit assumptions, the Film Unit constructed over a period of three years fourteen aptitude tests for which motion picture prints were available at the end of the war. They were, in most cases, conceived on the basis of the performances required of the pilot. Only two were based on an

*This chapter was drafted by Hibbard Lamkin, Alfred H. Shafer, and Robert M. Gagne. The tests themselves were constructed largely by cooperative effort and it would be difficult to assign individual credit and responsibility for the research. All the individuals listed in Chapter 1 shared in the insights, expedients, and prolonged labors which went into the making of these tests, with the addition of R. H. Henneman and S. R. Wallace, of the Perceptual Research Unit.

analysis of other performances—those of the bombardier—but since tests were to be validated empirically, it was possible that almost any test might prove valuable in selecting for almost any specialty.

Construction of these tests began early in the AAF Aviation Psychology Program. The potentialities of the motion picture had been realized at the outset in setting up the program of aptitude testing for aircrew candidates. Tests of three types were planned, printed tests, apparatus tests, and motion picture tests, each type being assumed to have its own unique advantages. By April 1942, a neucleus of what was to become the research organization in the Headquarters of the Training Command was on duty in the Office of the Air Surgeon in Washington, D. C., and it was determined that one of the responsibilities of this organization was to be the development of perceptual tests. In May 1942, the editor of this volume was directed to explore the agencies within the rapidly expanding Air Corps which might cooperate in film production for the purpose of constructing motion picture perceptual tests. Although training films were at that time being produced by the Signal Corps, this type of film production was scattered and relatively uncoordinated; it was therefore decided that the motion picture branch of the AAF experimental laboratory at Wright Field, Ohio, was the agency best adapted to produce the first motion picture tests. With the assistance of a film writer assigned by the Training Literature Section at AAF headquarters, the first scripts and specifications for motion picture tests were written and submitted to Wright Field early in June 1942. Because of reorganization of the motion picture branch, production of the films was delayed for some months and finally was sub-contracted to a commercial studio. The first three tests, involving animated photography and requiring judgments of velocity, were photographed under the supervision of the editor of this volume in October 1942, at the Metro-Goldwyn-Mayer studios in California. The tests were subsequently edited and completed as 16-mm. prints at Wright Field.

A great many difficulties had to be overcome in adapting the techniques of the sound motion picture to controlled psychological testing. Any novel use of a medium of presentation, or new combination of professional skills, probably involves the same difficulties. The motion picture had been used little if at all for psychological purposes and the motion picture technicians were unfamiliar with any other use of it than for storytelling or graphic reporting. The psychologist, therefore, spoke an unfamiliar language. Motion picture production had to be supervised at all stages by the psychologist in the making of tests, and the intricacies of motion picture production had to be learned.

As a result of this cooperation, however, a number of new possibilities both for psychology and for motion pictures have emerged.

The early aptitude tests constructed by the Perceptual Research Unit and the later ones completed by the Psychological Test Film Unit, which grew out of the former organization, will be described in the following pages. Accompanying each test are the results of administering them to samples of aviation cadets and aviation students. Only the tests themselves and the principal data obtained are reported, the research which preceded and accompanied them being omitted. In the case of a few tests in which this background research has general implications, such as those having to do with space perception, the results are more fully reported in another chapter.

Some of these tests were found to have predictive value for success in the various types of aircrew training, particularly pilot in training. None, however, appeared to be sufficiently outstanding in its contribution to justify incorporating it into the battery of tests actually used for recommending the appropriate specialty in which an aircrew candidate was to be trained. The validities of the tests, insofar as the data were obtainable, are given and will be summarized at the end. Many of the later and presumably more promising tests are lacking data from which to validate them because of the long delay between test administration and test validation.

APTITUDE TESTS CONSTRUCTED

Ability to Judge Motion and Locomotion

Estimation of Velocity Test CP205B–I. This test was constructed to measure the capacity to estimate and visualize the speed of an object moving at right angles to the line of regard. This ability may possibly be correlated with more general space-perceptual capacities to which names cannot be given. A test of such a function might also be related to success at gunnery.

The directions are given by subtitles with an accompanying voice on the sound track, together with examples and sample items. It is intended as a complex test. The judgment required is *where the object will be at a given instant.* Using animated photography, an airplane moves across a skyline background halfway across the screen and then disappears behind a cloud, as illustrated in Figure 5.1. The correctness with which the continued velocity of the imagined airplane is visualized is then measured by flashing a spot of light in the cloud ("anti-aircraft burst") and requiring a judgment as to the position of the imagined airplane in relation to the flash ("behind," "hit," "ahead") which is recorded on a 3-choice answer sheet. The judgment has been declared to be somewhat similar to that in "leading" the target in free gunnery.

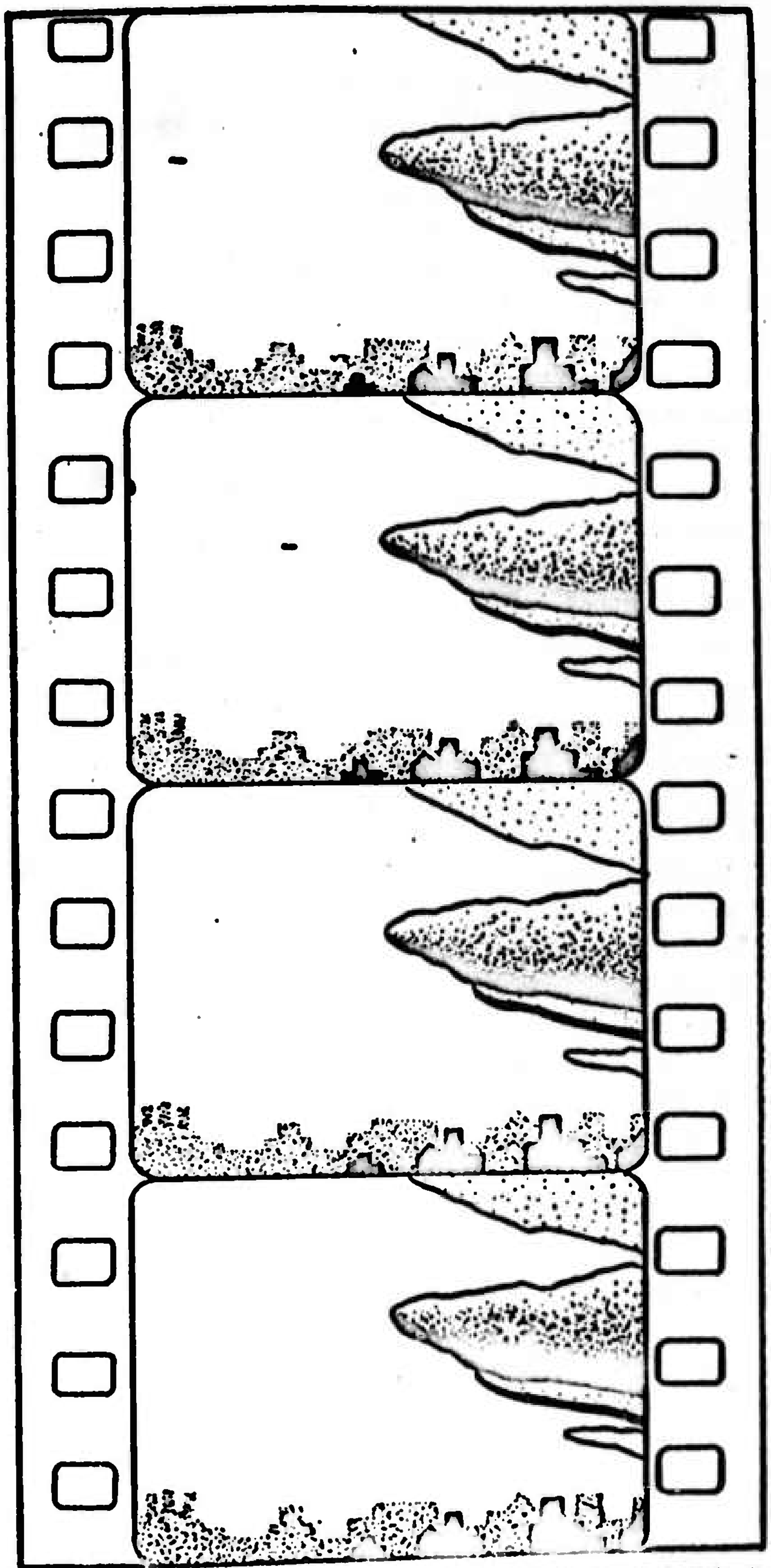


FIGURE 5.1.—Stages in a Sample Item of the Estimation of Velocity Test

Details of Development. Specifications for photographing the scenes, from which the trials of the test were to be selected, were drawn up in terms of a 9 by 12 inch picture frame. This is the size of the picture which is photographed, frame by frame, with the animation camera. Elements of the picture, on sheets of transparent celluloid, can be moved on a carriage between successive exposures of the camera. Motion was specified, on this scale, in sixty-fourths of an inch. Thus, a rate of sixteen-sixty-fourths of an inch, when shown at 24 frames per second, gives a velocity of 6 inches per second. In setting up artificial velocity stimuli with motion pictures, they should be visualized and specified by the test constructor on the 9 x 12 inch animation frame (or on a 9 x 12 foot projection screen, if desired) rather than in terms of the visual angle on the retina of an observer at a certain distance looking at a projected image of a certain size. The first method of specifying velocities is not only more convenient but more exact since, as was evident in chapter 4, the velocities perceived by the observers are those of the objective picture rather than those of their retinal images. An observer seated anywhere in an ordinary projection room sees substantially the same picture.

Since this test was photographed without preliminary experimentation, and since nothing existed in the experimental literature on velocity perception to indicate what discriminations of the type described were capable of being made by the average observer, a large number of scenes were shot, from which selections could be made of those at the proper level of difficulty after tryout and statistical analysis. Seven different velocities of the plane were photographed. Three positions of the flash in the cloud were shown. Most important, four degrees of deviation of the flash from the theoretically correct position of the unseen plane were shot. The deviations varied from three-sixteenths to twelve-sixteenths of an inch on the scale of the 9 x 12 inch picture. Eliminating some combinations of these variations, 96 shots were made which were estimated to vary from difficult to easy judgments. After viewing the projected shots, they were cut down to 64. These items were then spliced in random order as Form A of the test and given to 400 aviation cadets. An analysis of item difficulty was made and the items of too great or too little difficulty were eliminated. Forty-three different items were used to make up the final form of the test, without the necessity of repeating the same shot more than once. The original guesses as to appropriate velocities and deviations had proved to be very nearly correct. It is interesting to note that the velocity of an airplane, once seen, can be visualized with a fairly high degree of accuracy. On a 9 x 12 foot screen-picture, a deviation of the flash by one inch from the

position of the imagined airplane can usually be detected by the average observer.

Test Characteristics.

1. Administration and Scoring:

Form	Running time	No. trials	Scoring formula
	Minutes		
CP205A-I	30	64	R-W (March 1943)
CP205B-I	22	43	R-W (March 1943)

2. Reliability:

Form	Date	Place	Group	N	r	r_e	Type	Remarks
A	2/43	SAACC	Unclassified	407		.60	Hoyt	64 items; scored rights
A	2/43	SAACC	Unclassified	407		.65	Hoyt	43 items; scored rights
B	5/43	SAAAB	Unclassified 43-K	1323		.59	Hoyt	R-W.
B	12/43	MPEU#9	Unclassified 45-A	1000	.33	.50	Odd-even	R-W.
B	12/43	MPEU#9	Unclassified 45-A	1000	.38	.55	1st half, 2d Half	R-W.

3. Distribution Constants:

Form	Date	Place	Group	N	M	SD	Remarks
A	2/43	SAACC	Unclassified	407	42.8	5.5	R
A	3/43	SAAAB	Unclassified	392	20.15	10.89	R-W.
B	5/43	SAAAB	Unclassified 43-K	1323	15.07	7.55	R-W+10.
B	12/43	MPEU#9	Unclassified 45-A	1190	4.15	9.41	R-W.

4. Validity:

Form	Date	Place	Group	Type	N_t	P_g	M_g	M_e	SD	r_{bis}	r_{tet}	Remarks
A	2/43	SAACC	Unclassified 43-K	r_{bis}	250	.640	22.73	21.80	11.17	.05	.50	R-W.
B	5/43	SAAAB	Unclassified 43-K	r_{bis}	760	.726	25.08	25.03	7.47	.00	.02	Rights.

The criterion used was graduation-elimination from elementary pilot training.

5. Intercorrelations:

Correlated with—	Date	Place	Group	N	r
Identification of velocity CP205B-II	3/43	SAAAB	Unclassified	392	.12
Identification of velocity CP205B-II	12/43	MPEU#9	...do	1190	−.03
Estimation of relative velocity CP205B-III	3/43	SAAAB	...do	392	.02
Estimation of relative velocity CP205B-III	12/43	MPEU#9	...do	1190	.03
Plane formation CP805B	3/43	SAAAB	...do	392	.05

Identification of Velocities Test CP205B–II. This test was shot along with the one just described and was designed to measure the ability to discriminate visual velocities in a relatively "pure" form. The method of "absolute judgment" of velocities was employed, rather than successive comparison of paired velocities. Since the discrimination had to be based on an arbitrary scale, the test also involved memory for visual velocities. In gunnery, formation flying, and in other situations, the pilot must estimate the speed of other planes and his own and depend upon those estimations for appropriatie reactions. Presumably these estimations are related to simple visual judgments of velocity.

The moving stimulus employed was a relatively homogeneous field (clouds) rather than a moving object (plane). This prevented the subject from simply comparing the time intervals required for the object to cross the screen. Five increasing

velocities of the cloud background (with a physically stationary plane shown in the middle of the screen carrying the induced motion) were "taught" the subject and he was instructed to learn to identify them as A, B, C, D, or E. Speed A was reported to "look like" about 80 m.p.h. and Speed E, about 160 m.p.h. The five velocities increased by 20 percent at each step, which was estimated from available evidence to be neither too close to the differential threshold nor too far above it for testing purposes in the range of velocities employed. The 5 velocities were then presented 10 times apiece in a random order for identification by the subject. At intervals during the test, the "middle" speed, C, is presented for comparison in order to permit the testees to "anchor" their scale of judgment. This was all the more necessary since a negative time error would be expected to, and did to a considerable extent, affect the judgments.

Details of Development. The test was animated by photographing, frame by frame, a vague background, looking like fog, which moves behind the plane in the middle of the screen. The impression given is that the observer is flying parallel to the plane. The speeds shown vary from 6 inches per second to a little over 12 inches per second on the scale of the 9 x 12 inch picture. This represents successive increases of close to 20 percent. A series of 25 percent increases was also photographed and tried out, but proved to be too easy to discriminate. The normal differential threshold of velocity perception, for the situation described, seems to be such that the former series are identifiable on an absolute scale about 60 percent of the time by aviation students. The directions for the test, given by both titles and voice, are as follows:

This is a test of your ability to learn and identify different rates of speed. You will see a plane moving against a background of clouds, like this: (Speed C for 3 secs.). Five different velocities will be shown varying from a slow rate to a fast rate. The slowest rate will be indicated by the letter A and the fastest speed by the letter E. Study these scenes carefully, since you will later be asked to identify them by marking the appropriate letter on your answer sheet. Try to associate each scene with a definite feeling of its rate of speed so that you can recognize it later on. This is the slowest speed, A. (Speed A for 6 secs.) This is the fastest speed, E. (Speed E for 6 secs.) This is the middle speed, C. (Speed C for 6 secs.) You will see all five scenes, from the slowest to the fastest, shown in that order. (Scenes are shown.) These five scenes will now be shown to you in a mixed-up order, without any identifying letter. Each scene will be one trial, and trials will be numbered consecutively. At the end of each trial you will be allowed six seconds in which to identify the rate of speed. Mark the appropriate space on your answer sheet for the number of the trial. When you are not sure of your judgment, guess. Do not omit any trials. At regular intervals during the test there will be a pause and you will again be shown Speed C, the middle speed. This will enable you to check up on your judgments from time to time. All five tests of speed will appear

with about equal frequency. Now get ready for the test (pause). Here is Trial One. (Test begins.)—

Test Characteristics:

1. **Administration and Scoring:**

Form	Running Time	No. trials	Scoring Formula
	Minutes		
CP205A-II	17	50	R (March 1943).
CP205B-II	17	50	R (March 1943).

2. **Reliability:**

Form	Date	Place	Group	N	r	r_c	Type
A	2/43	SAACC	Unclassified	406	.52	.....	Hoyt.
B	5/43	SAAAB	Unclassified 43-K	1325	.60	.....	Hoyt.
B	12/43	MPEU No. 9.	Unclassified 45-A	1039	.44	.61	Odd-even.
B	12/43	MPEU No. 9.	...do	1000	.28	.44	1st half 2d half.

3. **Distribution Constants:**

Form	Date	Group	Place	N	M	SD
A	2/43	SAACC	Unclassified 43-K	406	31.90	4.4
A	3/43	SAAAB	Unclassified	392	31.70	4.9
B	5/43	SAAAB	Unclassified 43-K	1325	30.84	4.9
B	12/43	MPEU No. 9	Unclassified 45-A	1190	29.34	4.98

4. **Validity:**

Form	Date	Place	Group	Type	N_t	P_g	M_g	M_e	SD_t	r_{bis}	r_{tet}
A	2/43	SAACC	Unclassified 43-K	r_{bis}	250	.610	32.22	31.09	4.23	.16	.09
B	5/43	SAAAB	Unclassified 43-K	r_{bis}	767	.724	31.12	30.58	4.82	.07	.12

The criterion used was graduation-elimination from elementary pilot training.

5. **Intercorrelations:**

Correlated with—	Date	Place	Group	N	r
Estimation of velocity CP205A-1	3/43	SAAAB	Unclassified	392	.12
Estimation of velocity CP205A-1	12/43	MPEU No. 9	....do....	1193	-.03
Estimation of relative velocity CP205A-III	3/43	SAAAB	....do....	392	.15
Estimation of relative velocity CP205A-III	12/43	MPEU No. 9	....do....	1190	.18
Plane formation CP805B	3/43	SAAAB	....do....	392	.08

Estimation of Relative Velocities Test CP205B–III. This test was the third of the "Speed Estimation" tests and was designed to require a complex judgment of the *relation between two velocities,* measured by estimating the imagined point at which the faster of two moving spots (planes) would overtake the slower. The paths of the two motions are parallel. It was believed that the success or failure of an aviation cadet may depend in part upon his capacity to estimate relative or varied velocities for judgments required in training, combat flying and formation flying.

Two animated planes are seen against a skyline background, one overtaking the other. Before the overtaking point is reached, both planes disappear behind a cloud and a 5-point scale appears superimposed on the cloud, as illustrated in figure 5.2. The subject must "project" or imagine the two velocities and judge at which of the five points the two planes would coincide. This test consists of 50 trials, 10 at each of the five points, in a random order.

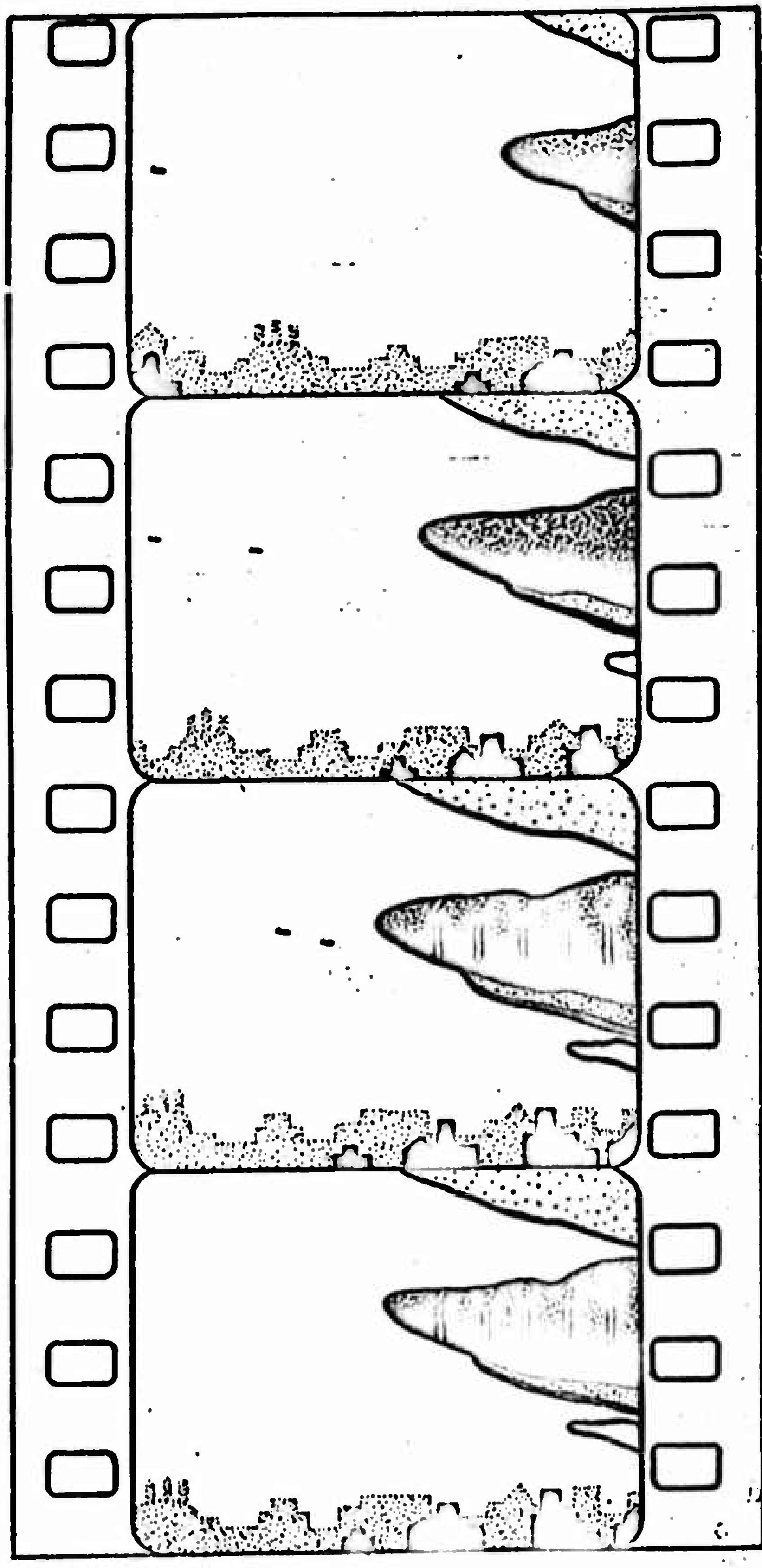


FIGURE 5.2.—Stages in a Sample Item of the Estimation of Relative Velocities Test.

Details of Development. The slower plane was animated at 2¼ inches per second and the faster plane at 3 inches per second in terms of the picture photographed. Five scenes were made showing the overtaking occurring at each of the five points on the scale and, since their level of difficulty was fairly satisfactory, the items of the test were made up from these five scenes only. The duplicate films representing these scenes were cut so that the beginning of the different items came at different points in the scenes as originally filmed, and the subjects were therefore unaware that some trials duplicated others. Because of the scale on which judgments of overtaking occurred, this test could be scored differentially instead of categorically by right or wrong judgments. The scoring formula can give part credit for judgments only one step removed from correct. This method was adopted. A statistical study showed that the weighted formula which maximizes validity is Rights (R) plus .85 of the adjacent, or one-step, Wrongs (W_1). Since all subjects necessarily finish the test and all were directed to attempt every item there is no need to deal with the problem of "omits."

Test Characteristics.

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula*
	Minutes		
CP205A-III	19	50	R (Mar. 1943).
CP205B-III	19	50	R+.85W_1 (Nov. 1943).

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	r_c	*Type*	*Remarks*
A	2/43	SAACC	Unclassified.	407		.60	Hoyt	R.
B	5/43	SAAAB	....do	454	.21	.34	1st half-2d half	R.
B	5/43	SAAAB	....do	454	.50	.67	Odd-even	R.
B	12/43	SAAAB	....do	503	.27	.43	1st half-2d half	R+ .85W_1
B	12/43	MPEU No. 9.	....do	1,000	.35	.52	1st half-2d half	R+ .85W_1
B	12/43	MPEU No. 9.	....do	1,000	.50	.67	Odd-even	R+ .85W_1

3. Distribution Constants:

Form	*Date*	*Place*	*Group*	*N*	*M*	*SD*	*Remarks*
A	2/43	SAACC	Unclassified	406	28.6	5.3	R.
B	3/43	SAAAB	....do	460	29.18	4.66	R.
B	5/43	SAAAB	Unclassified 45 K	1329	29.31	5.25	R.
B	12/43	SAACC	Unclassified 44 G, H, I, J	1047	44.71	2.56	R + .85W_1.
B	12/43	SAACC	....do	1047	28.76	5.13	R.
B	12/43	NAAC	Unclassified	1068	29.92	5.26	R.
B	12/43	MPEU#9	Unclassified 45 A	1190	28.83	5.48	R.

4. Validity:

Form	*Date*	*Place*	*Group*	*Type*	N_t	p_s	M_s	M_u	SD_t	r_{bis}	σr_{bis}	r_{bis}	*Remarks*
A	2/43	SAACC	Unclassified	r_{bis}	250	.640	29.27	27.47	5.39	.21		.07	Rights
B	5/43	SAAAB	Unclassified 43-K	r_{bis}	768	.718	29.54	28.28	5.32	.14		.04	Rights
B	12/43	SAAAB	Unclassified 44-K	r_{bis}	396	.506	44.61	44.49	2.69	.03	.03	.01	R+ .85W_1
B	12/43	SAACC	Unclassified 44G,K	r_{bis}	1047	.599	44.82	44.22	2.56	.13	.15	.05	Rights
B	12/43	SAACC	Unclassified	r_{bis}	1047	.599	....	....	....	.18		.05	R+ .85W_1

The criterion used was graduation-elimination from elementary pilot training.

5. **Intercorrelations:**

Correlated with—	*Date*	*Place*	*Group*	*N*	*r*
Est. of velocities CP205B-I	3/43	SAAAB	Unclassified	392	.02
Est. of velocities CP205B-I	12/43	MPEU No. 9	...do	1,190	.03
Ident. of velocities CP205B-II	3/43	SAAAB	...do	392	.15
Ident. of velocities CP205B-II	12/43	MPEU No. 9	...do	1,190	.18
Plane formation CP805	3/43	SAAAB	...do	392	.10
Rudder control CM120A	1/44	SAAAB	...do	431	—.05
Finger dexterity CM116A	1/44	SAAAB	...do	....	—.02
Rotary pursuit CM803A	1/44	SAAAB	...do	....	—.01
Complex coordination CM701A	1/44	SAAAB	...do	....	.00
Discr. react. time CP611D	1/44	SAAAB	...do	....	.03
Two-hand coord. CM101A	1/44	SAAAB	...do	....	.04

Landing Judgment Test CP505E. This test was designed to measure the ability to learn certain spatial discriminations believed to be required for successfully landing an airplane. The specific judgment required is the direction of one's movement in an approach glide toward a landing strip, and involves the use of the "expansion cue" and other cues. Basically it is a test for space perception, using the landing situation, in which the visual variables ("cues") are experimentally controlled. The ability to utilize the pattern of apparent movement of the ground in order to judge the point at which the plane is aiming is considered important in performing a landing. A description of the expansion cue as an important stimulus for space perception is given in Chapter 10. In addition, the more general ability measured by this test, that of learning rapidly to make spatial discriminations, seems to be of considerable importance to the pilot, whose job makes necessary a large number of such perceptual judgments.

A view of a runway and part of a surrounding airfield is shown on the screen, as it appears during an approach glide. Five spots, lettered A, B, C, D, and E are located on the closest third of the runway, and each trial consists of a ten-second glide toward one of these spots. The candidate is required to judge at which of the five spots on the runway his glide is aimed, and to record his choice on a standard answer sheet. Optically, the spot being approached remains stationary, while the rest of the visual field expands or enlarges outward from this central point. This factor provides a visual cue for judging the direction of movement. Three photographs of frames from the top, middle and end of a glide are presented in figure 5.3. In the approach glide from which this illustration was obtained, the expansion was centered about spot D toward which the glide was aimed.

Details of Development. Preliminary forms of the test (CP505A, B, C), consisting of series of experimental shots obtained with a camera mounted in a real plane during a landing glide, were found to be impractical. The set of shots (Form D) taken at the AAF First Motion Picture Unit was based on the requirement of having five lettered aiming points on the runway, with five corresponding scenes of an approach glide, each taken at three different speeds of approach. In order that the

glides could be made in a perfectly straight line with stable motion, a model airfield was employed (scale 1 to 48) and the camera was moved on a track set at an angle of 30° to the horizontal. This track could be put in five positions, without varying the angle, in such a way that the path of the camera was aimed at each of the five spots. The camera was moved at a constant speed of 1.2 feet per second through a distance corresponding to one-quarter of a mile (1,440 ft.), down toward the runway. The glide represented was steeper and slower than that obtained with a military airplane, but this was believed to be essential for the purpose of obtaining adequate degrees of difficulty and variability in test items. Each of the five glides was photographed at three different shutter speeds of the camera (36, 27, and 18 frames/sec.) so as to produce apparent speeds of 26, 35, and 52 miles per hour. The lens used gave a field of view 35° in width and 26° in height. The 15 basic scenes obtained in this manner constituted the preliminary form (D) of the test, and were used with modifications to make up its final form.

The test was now given for the purpose of investigating two hypotheses, before the final form was constructed: (a) that there would be a tendency for the center of the screen, rather than the center of the expansion pattern, to be judged as the aiming point; and (b) that verbal instructions on the specific visual cue involved would hasten the learning of the spatial discriminations required.

It was found that judgments made by using the spatial position of the aiming point in the center of the screen did actually occur in the case of most observers. The use of this artifactual "screen cue" rather than the expansion pattern was avoided by offsetting the center of the scene with respect to the frame. In this way the aiming point was always above or below the center of the screen. The aiming points of 8 of the 15 basic scenes were offset upward and 7 were offset downward from the center of the screen in a random order. Fewer than 5 percent of a group of cadets presented the off-centered scenes reported that they had attempted to use the centering cue. The results of the experiment on this cue may be found in more detail in chapter 9.

Considerable effort was directed toward obtaining a script which would give an accurate description of the expansion pattern and instruction in its use to find the aiming point of the approach glide. A script was finally developed which when given prior to the test, had the effect of a training or instructional period on the perceptual basis of landings.

Evidence was obtained (and is presented in chapter 9) that this preliminary instruction does actually yield a substantial amount of learning. One group of students were given 45 trials with perfectly clear directions as to what judgment they were

required to make, but no instruction in how to make it. There was no improvement in the score in the last 15 trials over the first 15, and scores were low. Another group was given instruction and practice during the middle 15 trials together with the opportunity to correct their errors. This group improved to a marked degree in the last 15 trials and scored high.

Both the factors of verbal instruction and off-centering were taken into account in the construction of the final form of the test. The first 10 minutes of the film are used to give instruction on how to judge the point of aim during an approach. The scene showing the field and runway is shown repeatedly on the screen, while a commentary and titles give instruction in observation of the expansion pattern cue. The subject practices making the judgment, and his errors, if any, are corrected during 10 practice trials. By means of optical printing, the 60 test items and 10 practice items of the final test were all off-centered either upward or downward, so that the correct spot never appeared in the center of the screen. The length of each item is 10 seconds, this length having been shown, in preliminary testing, to yield about the right degree of difficulty. The items in the final form were arranged in random order with respect to the five aiming points, the three-speeds, and the two types of off-centering. The nearness of the camera to the ground at the end of the 10-second approach (determined by cutting the total scene) is decreased in the last 30 trials so that they are made more difficult than the first 30.

Test Characteristics.

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula*
	Minutes		
CP505D	8	15	R.
CP505E	28	60	R + .5W_1 (one-step wrongs).

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	r_e	*Type*	*Remarks*
D	4/44	SAAAB	Unclassified preflight.	66	.34	.68	Test-retest	r_e for 60 items (repeated immediately).

3. Distribution Constants:

Form	*Date*	*Place*	*Group*	*N*	*M*	*SD*	*Remarks*
D	5/44	SAAAB	Unclassified preflight.	300	9.49	2.21	Form D (15 items).

4. Validity:

No data are available for inclusion here.

Ability to Judge Distance

Distance Estimation Test CP212A. This test was designed to measure the ability to make spatial discriminations based upon the perception of distance. It is generally agreed that such judgments constitute an important part of the pilot's task. Existing

tests designed to measure the perception of distance or depth involve the presentation of stimulus objects placed at a distance of 20 feet or less from the individual being tested. Yet judgments of distances of such small magnitude must be required of the pilot very infrequently insofar as the essential aspects of his job are concerned. He has to deal, in landing and take-offs, for example, with distances of hundreds of yards. It is possible that the cues which chiefly determine the perception of short distances such as those represented in the existing tests, namely the binocular cues of retinal disparity and convergence and the monocular cue of accommodation, diminish with increasing distance and have finally disappeared at these greater distances. Therefore it may be argued that existing tests do not measure the function of distance perception as it is required in the pilot's task. The present test was designed to involve the perception of these large distances. It is based on the hypothesis that the visual cues for the perception of large distances are such that they can be reproduced by photographic means. This hypothesis is examined in detail in chapter 9.

The test consists of a series of twenty 9 x 12 inch glossy photographic prints. In the foreground of each photograph may be seen a series of fifteen numbered thin white stakes, arranged in order from shortest to highest. They are arranged along a portion of a large arc, and are equidistant from each other, except for a somewhat larger gap in the center, between stakes 8 and 9. Through this central gap, off in the distance, may be seen another white stake (the test object) against a background of small hills. This arrangement is illustrated in figure 5.4. The terrain in the photograph is a perfectly level dirt field, which eventually joins hills and shrubbery on both sides and in the far background at a distance of about one mile. It was selected as one in which good cues for distance perception were present and could be adequately represented in a photograph. Four different heights of the test object are shown at five different distances in successive photographs. The candidate's task is to match the real height of the test object in the distance with the height of one of the 15 standards in the foreground. The test itself is preceded by instructions which emphasize that the natural judgment of real size is desired rather than size on the photograph, and two example items aid in establishing this attitude on the part of the candidate.

An experiment was performed to determine the degree of correspondence between judgments made using the photographs and judgments obtained in the actual field situation represented by the photographs. The results are reported in detail in chapter 9. It may be briefly stated here that a comparison of the judgments of thirteen subjects in the two situations reveals a high degree

of relationship between them. Photographs can be used to represent distances as far as 1/4 mile. The experiment therefore provided evidence that genuine distance perception could be measured with only photographic representation of the distance.

The objects represented in the photographs are as follows: (a) fifteen "standard" stakes, ranging in height from 27 inches to 83 inches, differing from each other by 4 inches, and of widths which varied in random fashion from two to four inches; (b) 4 test objects, of similar widths, of 63 inches, 67 inches, 71 inches, and 75 inches in height. The actual distances from the camera at which each test object was photographed were 28, 56, 112, 224, and 448 yards.

Two sets of these 20 photographs were each arranged in random order and given as a test of 40 items, with preceding instructions and examples, individually to 50 subjects. As the data presented in a later paragraph indicate, scoring this test by the formula "Rights" yielded scores which were far too low and resulted also in low reliability. As in a number of other psychological tests where the alternatives represent scale deviations from correct, there is good reason to credit wrong responses in proportion to the degree of correctness they represent. Consequently a scoring formula which weighted rights 3, one-step wrongs 2, two-step wrongs 1, and three- or more-step wrongs 0, was tried on this test and found to be satisfactory.

Test Characteristics

1. Administration and Scoring:

Form	*Administration time*	*No. trials*	*Scoring formula*
CP212A....................	18 min..................	40	$3R+2W_1+W_2$ (W_1=one-step wrongs; W_2=two-step wrongs).

2. Reliability:

Date	*Place*	*Group*	*N*	M_a	M_b	SD_a	SD_b	*r*	r_c	*Type*	*Remarks*
6/45	SAAAB	Returnee air crew	50	3.04	3.20	2.05	2.04	.40	.57	1st half-2d half	R.
6/45	SAAAB	....do....	50	24.24	25.08	[illegible]	8.08	.66	.79	1st half-2d half	$3R+2W_1+W_2$

3. Distribution Constants:

Date	*Place*	*Group*	*N*	*M*	*SD*	*Remarks*
6/45	SAAAB	Returnee air crew...	50	6.42	3.42	Scored R.
6/45	SAAAB	....do..............	50	49.26	16.4	Scored $3R+2W_1+W_2$.

4. Validity:

No data are available as the test was never administered to a sample of aircrew candidates.

5. Intercorrelations:

None are available.

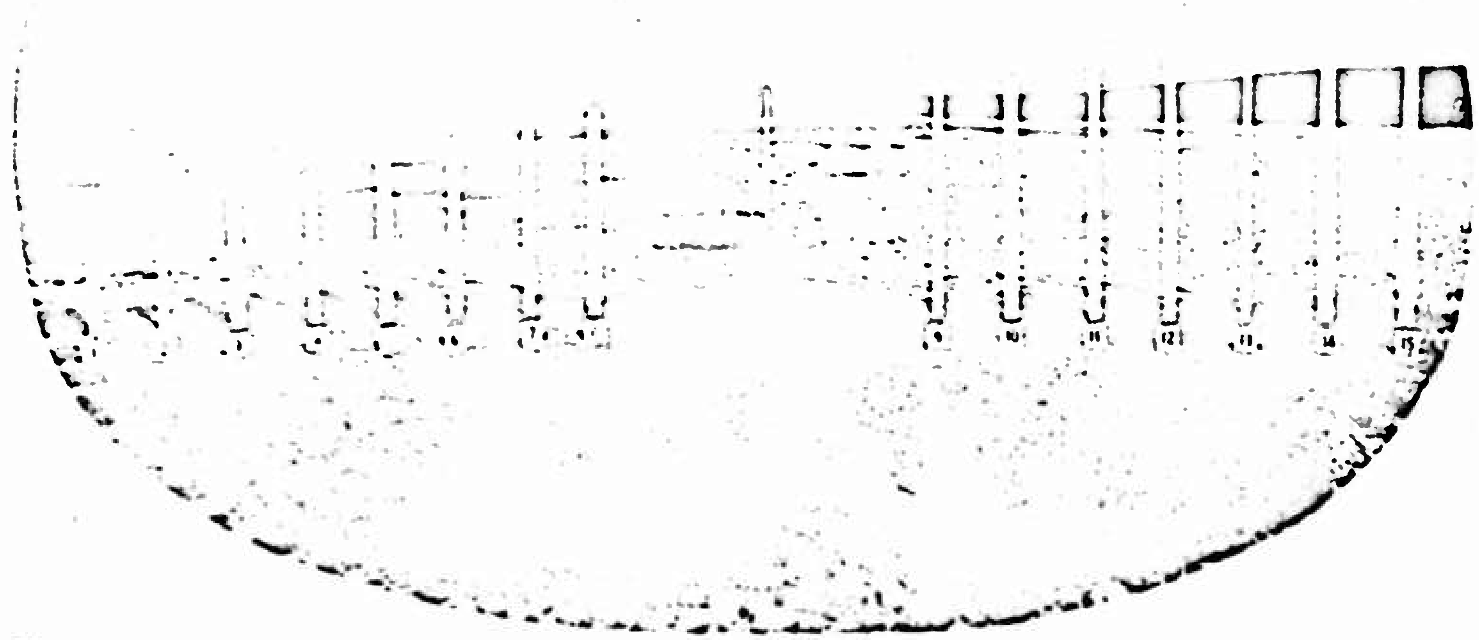


FIGURE 5.4.—Sample Item of the Distance Estimation Test, Form A. · In this sample the test object is placed at a distance of 28 yards and matches in real height standard No. 11.

Ability to Maintain Orientation in Space

The ability to maintain orientation in the air is obviously even more important than it is on the ground. Strong individual differences seem to be present for the ability. Artificial devices such as maps, compasses, signs, arrows, and the like are necessary for long continued orientation, but a flier should have what is loosely called a "sense" of orientation, i. e., of where he is headed and where he is in *visible* space. The ability to infer heading and location from artificial aids is a secondary and more intellectual ability. The former is presumably a spatial ability. After a considerable degree of preliminary research, two tests were constructed in this field.

Flying Orientation Test CP107A. This test was designed to measure the ability to maintain directional orientation when flying, together with the allied ability to visualize a flight path already flown. It is believed that a number of requirements of successful flying may involve these abilities. Other tests designed to measure orientation in flight suffer from the limitation that they measure, in part at least, certain abilities such as form perception and that certain intellectual factors enter into the situation, i. e., correct answers can be obtained by treating the items of the test in a "problem-solving" manner. This test, by utilizing a motion picture representation of the ground during flight, should secure a type of response similar to the actual maintenance of orientation in flight as distinct from orientation implemented by artificial aids such as the compass used in conjunction with a map.

The candidate sees on the motion picture screen a view as if he were looking straight down from a plane flying high above the ground. The appearance of looking down from within the bomb-bay of a plane is given by a representation of open bomb-bay doors along the sides of the screen, as illustrated in figure 5.5. During each test trial the ground is seen moving as it would appear in a plane whose course is composed of several straight legs of constant length, with intervening turns of 90° to the right or left. The task of the observer is to maintain his orientation, by watching the ground, so that at the end of a flight involving several turns he can indicate the direction of the starting point from his present position. The vertical view of the ground is employed in this, rather than any other, because the experiment showed that only in this way could "artifactual" cues to orientation and memory for landmarks be eliminated from the judgments.

At the end of each flight pattern, the view of the terrain fades into a circle with eight arrows, lettered from A to H, similar to a compass rose. Arrow A always points straight ahead (toward the top of the screen), arrow B is 45° to the right from arrow A, arrow C is 90° to the right, etc. Hence eight directions are pos-

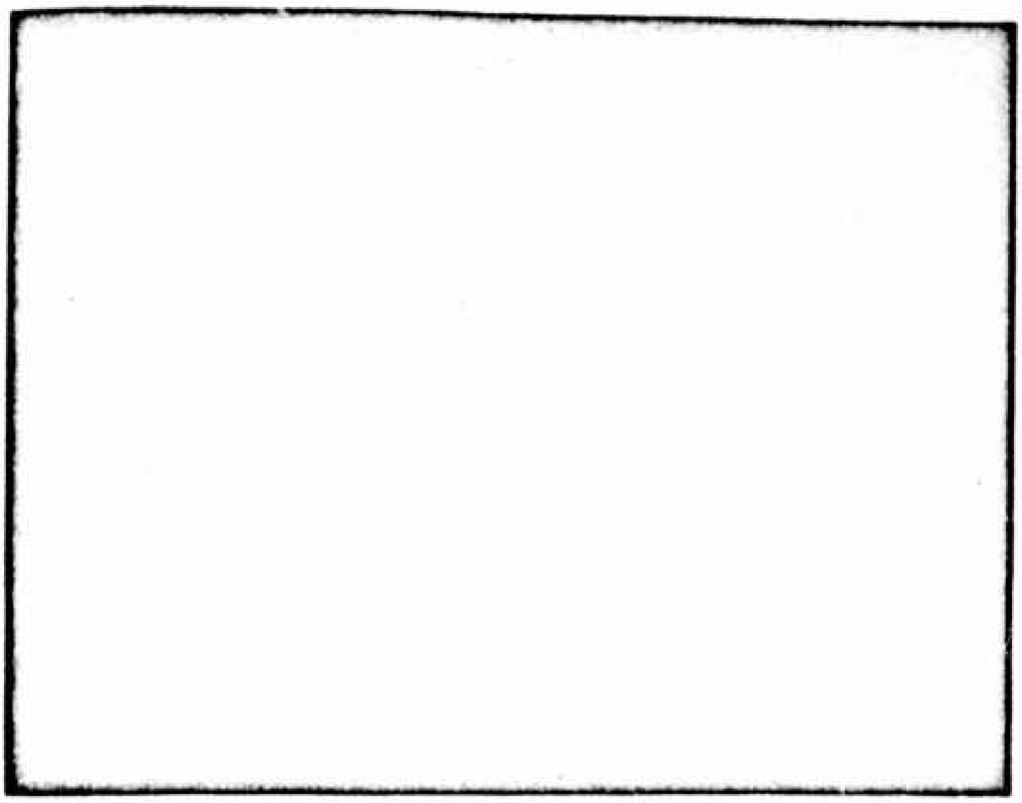


FIGURE 5.5.—View of Terrain Through Bomb-Bay Doors.

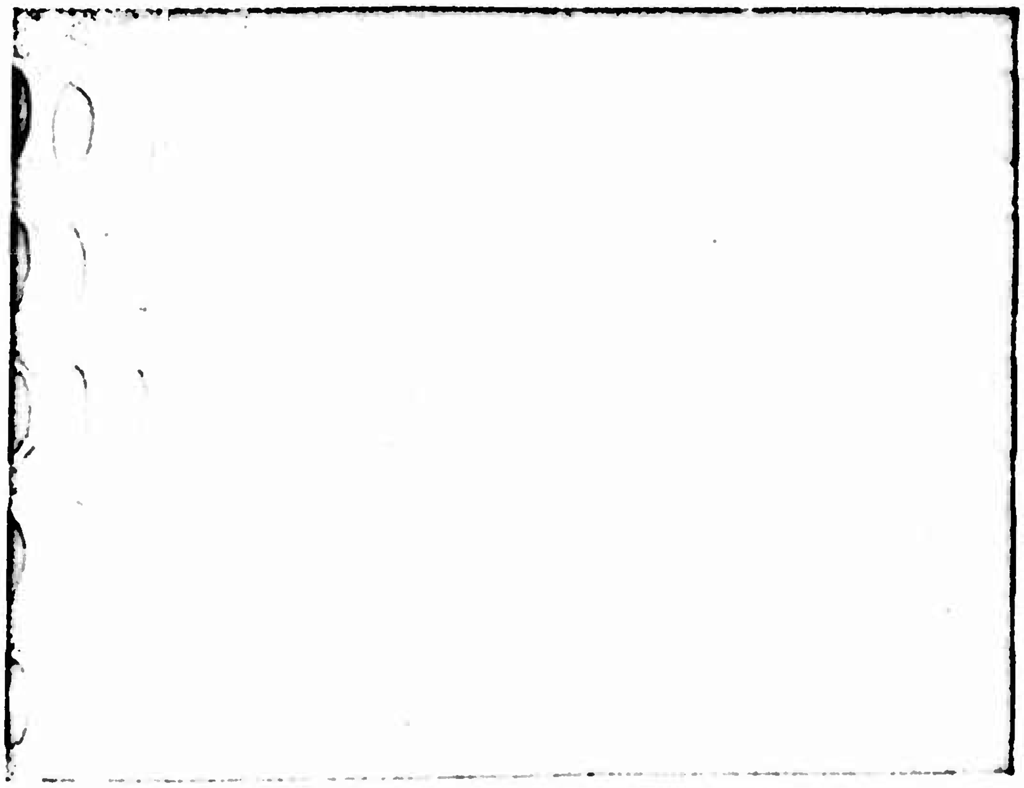


FIGURE 5.6.—Over-All View of Path Flown Over Terrain. This view represents a sample path as it would be seen from a great altitude. Note that the path follows the lines of a checkerboard pattern.

sible, each differing from the next by 45°. The subjects have to indicate the direction of the starting point by choosing the one of the eight lettered arrows which points toward it, entering the letter of this arrow on an I.B.M. answer sheet. The directional response is made from the point of view of the orientation of the plane and the observer rather than of a compass direction on a map.

In the photography of the test, an enlarged aerial photograph was used as a substitute for the ground. The motion picture camera moved over the aerial photograph in the same way that a plane with an aerial camera would fly over real terrain. The path followed by the camera consisted of from four to seven legs of constant length with intervening 90° turns, either to the right or left. This is illustrated in figure 5.6. On the screen, the movement of the terrain is always from the top of the screen toward the bottom, i. e., when the plane makes a 90° turn, the terrain swings around 90° and then continues moving from top to the bottom of the screen. The test consists of 25 "flights" of about 30 seconds duration, with a total running time of approximately 30 minutes.

In order to maintain orientation to his starting point, since the starting point is visible only at the beginning of each flight, the observer must not only keep track of its direction but also the distance he is from it. This makes it necessary for him to visualize the path he has taken over the ground, from the start to his present location. This ability consists of two general performances, probably interrelated: (a) The subject must observe the ground during a turn, interpreting its movement correctly as to whether the plane is turning left or right, and visualizing the relation of the starting point to the plane during the turn; (b) he must be able to visualize the total path he has followed over the ground. The directions given the examinee were presented fully in chapter 3.

Details of Development. Experimental work with preliminary apparatus to determine items and their levels of difficulty was carried out. Suitable aerial photographs were obtained, accessory apparatus necessary for photography was constructed, and preliminary test photography completed. Final photography, using a 16 mm. camera, was accomplished by the Test Film Unit with its own facilities. A number of 16 mm. prints of the test were made on silent film. The directions are given fully by titles. They may also be read aloud by the test administrator for extra emphasis as they appear.

Administration and Scoring:

Form	Running time	No. trials	Scoring formula
CP107A	24 minutes	25	R

No other statistical data are available on this test as it was completed in October 1945.

Landing Orientation Test CP106A. The purpose of this test was to measure the ability to discriminate, learn, and remember the features of the ground that serve as cues for spatial orientation in the traffic pattern. In flying the traffic pattern preparatory to landing, the pilot must make appropriate flying responses based upon his position in space and hence upon the appearance of the ground corresponding to those positions. It is judged that the ability to learn to discriminate various positions in the traffic pattern from the appearance of the ground can be measured adequately by a motion picture test.

In this test the observers are shown a view of a landing field as it appears from a plane flying a modification of a standard traffic pattern. The flight path can best be described as a slow wheeling turn from a point half-way through the down-wind leg to the end of the base leg of the traffic pattern. During the flight the camera is oriented so that it is always pointed at a specific area of the landing field. The reason for using this path and type of photography is so that a minimum of translatory movement would appear on the screen to provide "screen" cues for learning and remembering the appearance of the ground at the crucial moments.

An item in the test consists of two short sections of the flight course, the second run being a repetition of the first. In the first run the subjects are told that they are to watch the ground so that they can note their position when the narrator's voice says, "Now!", since they must be able to recognize and designate that position during a repetition of the same run. At a selected point after the run begins, the narrator's voice says, "Ready" (pause) "Now!". A few seconds later the scene ends and then the scene begins in approximately (but not exactly) the same place as before. As the "plane" approaches the spot where the "now" signal was given, the narrator announces, "Ready" (pause) "A—B—C—D—E—Stop" giving the letters at two-second intervals. The subjects, as directed in the preparatory instructions, will choose and record on their answer sheets the letter which preceded the interval in which the critical position fell.

In developing this test, a 97-foot section of the aerial film taken by First Motion Picture Unit cameramen was chosen for the construction of a preliminary form of the test. This section, taken from a blimp over an air field at about 800 feet at 60-70 m.p.h., shows a wide wheeling turn along the side and end of the runway. Fourteen copies of this section were secured to use in constructing an experimental form of the test. A method of marking the film so that an operator can watch it going through the projector and

time his narration and signals was devised. Instructions were written, and the test was administered to subjects to determine the optimal spacing of the letters A, B, C, D, and E, the difficulty level of the items, and the clarity of the instructions. The revised instructions exist in the form of a synchronized script to be read by the test administrator. The film exists as a spliced 35 mm. print, suitable for projection as a demonstration or experimental test, but without the full set of items necessary for a completed test. Specifications are written for the complete test of 50 items, lasting 35 minutes, to be scored "rights."

Perception of Slight Movement

The duties of a bombardier require the ability to "synchronize" the bombsight promptly and accurately in the early minutes of a bombing run over the target. A pair of crosshairs must be centered on the target, and adjusted for "rate" and "course" (i.e., in two dimensions) so that they do not drift away from the target in either of these dimensions. The compensating motion of sight must be so adjusted that it exactly "kills" the relative motion of the plane to the ground. This description, although oversimplified, makes clear that the ability is required to react promptly to barely detectable drift of the target relative to the crosshairs, and also to barely detectable inclinations of such drift toward or away from the vertical and horizontal axis formed by the crosshairs. Individual differences among observers exist in just noticeable movement; likewise the ability to perceive drift undoubtedly differs among bombardier students. Two tests of this ability were devised, after a theoretical analysis of the perceptual task of the bombardier, based on a visit to a bombardier school and interviews with instructors.

Minimal Movement Test CP213C. This test was designed to measure the ability to detect barely visible movement of an object and to determine the direction of this movement.

In the test, a black spot appears on the screen against a gray background in one of the four quadrants of the field of a bombsight, schematically represented. It appears in different quadrants for different items. In figure 5.7 the spot appears in the "Northeast" (NE) quadrant. The spot is stationary for the initial 2 seconds of each trial. It may then either remain stationary or it may move very slightly in one of four directions: either up, down, right, or left, (U, D, R, or L). The task is to indicate whether the spot remains stationary or, if it moves, to record the direction of motion. The items vary in respect to the speed of motion of the spot and the length of time during which the movement appears on the screen (2-3 secs.).

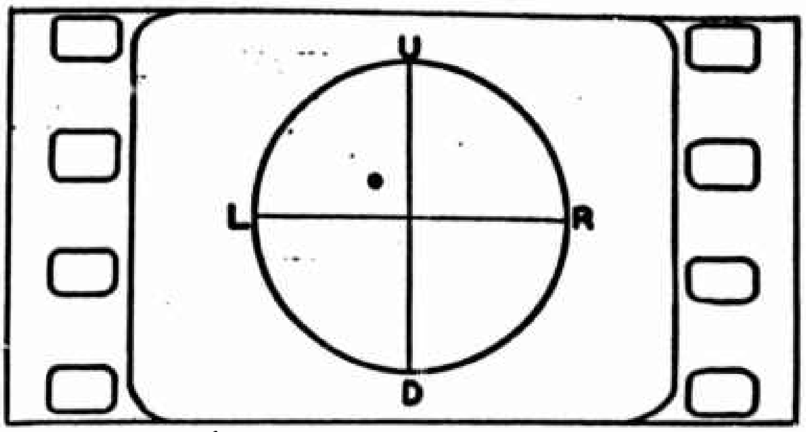


Figure 5.7.—Appearance of Sample Item of Minimal Movement Test.

Details of Development. In the preliminary form of this test, Form B, the black spot moved at two different rates of speed. Expressed in terms of a 9 x 12 inch animation frame, these rates were .047 and .031 inches per second. When projected for experimental purposes at silent speed, (16 frames/sec.) these speeds became .031 and .021 inches per second respectively. The test with these speeds represented a relatively easy task. Form C of the test was designed to overcome this difficulty and to be run at sound speed (24 frames/sec.). The four velocities of the spot employed in this form were .031, .024, .019, and .016 inches per second expressed in terms of a 9 x 12 inch frame. (These may be specified as 1/750, 1/1000, 1/1250, and 1/1500 inches per frame, respectively.) Other variables which were considered in designing Form C were the following: Quadrant in which spot appears, direction of motion of spot, and length of item. The test was constructed as follows:

Trials	*Speed*	*Length*	*Quadrants*	*Directions*
		Seconds		
1-3 (practice)	1/750	3	NE, SW, SE	U, R, S.
4-13	1/750	3	5 NW, 5 SE	2 ea UDRLS.
14-23	1/750	2	5 NE, 5 SW	2 ea UDRLS.
24-33	1/1000	3	5 NW, 5 SE	2 ea UDRLS.
34-43	1/1000	2	5 NE, 5 SW	2 ea UDRLS.
44-53	1/1250	3	5 NW, 5 SE	2 ea UDRLS.
54-63	1/1250	2	5 NW, 5 SW	2 ea UDRLS.
64-73	1/1250	3	5 NW, 5 SE	2 ea UDRLS.
74-83	1/1250	2	5 NE, 5 SW	2 ea UDRLS.

The directions were randomized throughout each set of ten trials, and included 2 *still* (S) items for each ten trials. The quadrants also occur in random order within each set of 10 trials.

Test Characteristics

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula*
	Minutes		
CP213B	15	63	R.
CP213C	20	83	R.

2. Reliability:

Form	Date	Place	Group	N	r	r_c	Type	Remarks
B....	1/44	SAAAB	Unclassified preflight	165	.63	.77	Odd-even........	Slow projection. .....
B....	1/44	SAAAB	....do.............	164	.54	.70	1st half-2d half..	Slow projection. .....
B....	6/44	SAAAB	....do.............	222	.63	.77	Odd-even........	3 sec. items only (20).
B....	6/44	SAAAB	....do.............	222	.52	.69	Odd-even........	1 sec. items only (20).

3. Distribution Constants:

Form	Date	Place	Group	N	M	SD	Remarks
B	11/43	SAAAB	Unclassified preflight	370	53.94	5.76	Test too easy.
B	1/43	SAAAB	....do.............	165	47.44	5.85	Slow projection.

4. Validity:

No data on the validity of this test were obtained, owing largely to the fact that no satisfactory criterion against which to validate it existed. The Psychological Research Project (Bombardier) undertook to validate the test against synchronization scores on a bombing trainer. The data were not available for inclusion here.

5. Intercorrelations:

Correlated with—	Date	Place	Group	N	r
Drift direction (Form B)..	2/44	SAAAB	Unclassified preflight.	164	.21
Pilot stanine	2/44	SAAAB	....do..............	148	.19
Bomb. stanine	2/44	SAAAB	....do..............	148	.22

It may be noted that the function tested has a low relationship with that measured by the Drift Direction Test, next to be described.

Drift Direction Test CP 221B. This test was designed to measure the ability to detect the drift of a moving spot to one side or the other of the main direction in which it moves. It is essential for a bombardier, as the target is approached, to be able to detect any slight deviation from the correct setting of the drift-knob, i. e., any tendency of the target to drift to one side of the crosshairs.

A black spot appears against a gray background in one of the four quadrants of a circle with perpendicular crosslines, like the field of a bombsight. The spot is stationary for the initial 2 seconds of each item. The spot then moves, slowly but noticeably, either up, down, to the right, or to the left, depending upon which of the four quadrants it is located in. Figure 5.8 illustrates the movement of the spot when located in the NE quadrant. In some items the spot moves in a perfectly straight line parallel to one of the crosslines of the circle; in other items it drifts toward or away from the crossline. Drifts of 3° of inclination were determined empirically to be just noticeable in this situation for a sample of aviation students, and this inclination is used in the final form of the test. Subjects must detect this drift, if it exists, and also its direction, and respond accordingly in one of three categories (e. g. to the right; to the left; or straight). The items vary with respect to the quadrant in which the spot appears, the direc-

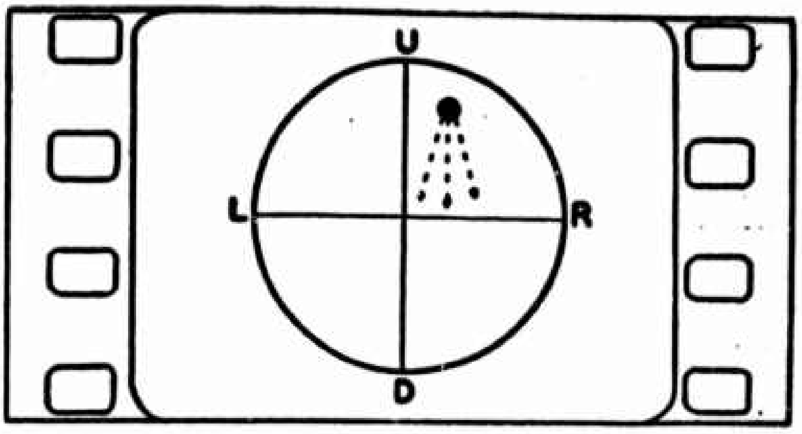


FIGURE 5.8.—Types of Movement of the Spot in the Drift Direction Test. The slant of the paths represented is exaggerated. The spot shown is in the NE quadrant; therefore the judgments might be L, R or S.

tion of drift, and the length of time during which the spot appears on the screen.

Details of Development. Form A of this test contained two different degrees of drift, 3° and 4°, the speed of movement of the spot being about a half-inch per second for a 9 x 12 inch picture. Experimental administration of this test showed it to be too easy. The difficulty level was shown to be approximately right when the 4° items were eliminated from the test and when the motion of the spot was made slower. Form B of the test was constructed in order to achieve the proper difficulty level. The items all have a 3° drift, and the spot moves at a rate of 1/64 inches per frame or .375 inches per second (at 24 frames/sec.) on a 9 x 12 inch picture. The duration of movement of the spot is varied systematically, from 2 to 6 seconds. Details of the construction of Form B are shown in the following table:

Trials	*Length* *Seconds*	*Quadrants*	*Directions*
1-6 (practice)	6	NE, NE, NE, SW, SE, SW	2 ea U, D, L, R, 4S
7-18	6	3 ea NE, NW, SE, SW	2 ea U, D, L, R, 4S
19-30	5	3 ea NE, NW, SE, SW	2 ea U, D, L, R, 4S
31-42	4	3 ea NE, NW, SE, SW	2 ea U, D, L, R, 4S
43-54	3	3 ea NE, NW, SE, SW	2 ea U, D, L, R, 4S
55-66	2	3 ea NE, NW, SE, SW	2 ea U, D, L, R, 4S

The directions and quadrants are randomized throughout each set of 12 trials. Four "straight" items are included in each set of 12, along with two of each other direction. "Straight" items are more difficult than items with an inclination "up" or "down," "right" or "left," as the case may be.

Test Characteristics

1. Administration and Scoring:

Form	*Running time* *Minutes*	*No. trials*	*Scoring formula*
CP221A	22	90	R.
CP221B	19	66	R.

2. Reliability:

Form	Date	Place	Group	N	r	rc	Type	Remarks
A....	1/44	SAAAB	Unclassified preflight	164	.42	.59	Odd-even	Slow projection
A....	1/44	SAAAB	....do.............	164	.45	.62	1st half-2d half	....do........

3. Distribution Constants:

Form	Date	Place	Group	N	M	SD	Remarks
A.....	11/43	SAAAB..	Unclassified preflight	226	55.54	8.07	Test too easy.
A.....	1/43	SAAAB..	....do..............	164	55.86	6.71	Slow projection.
A.....	1/43	SAAAB..	....do..............	164	39.92	5.20	66 items (3° drift only).

4. Validity:

No data are available owing to the absence of a satisfactory criterion. The results of a validity study against bombing trainer scores previously referred to could not be obtained for inclusion here.

5. Intercorrelations:

Correlated with—	Date	Place	Group	N	r
Minimal movement CP213B..	2/44	SAAAB	Unclassified preflight.	164	.21
Bombardier stanine...........	2/44	SAAAB	do..................	148	.16
Pilot stanine (cr. excl.)......	2/44	SAAAB	do..................	148	.24

Attention may be called again to the fact that there is apparently a low relationship (.21) between the ability to detect *motion* of the spot relative to the crosshairs and the ability to detect the *inclination* of the motion toward or away from the axis to which it is approximately parallel.

Multiple Perception

Flexibility of Attention Test CP411E. This test was designed to measure the ability of an aircrew candidate to distribute his attention over a wide range of stimuli. The pilot must take account of a number of different events occurring in different parts of the perceptual field at the same time. He must do so without becoming confused, and he must react appropriately to the different events. Such an ability has no close parallel in any of the laboratory investigations of attention, and it is probably highly complex. It was thought that it would be possible to test this aptitude by means of a motion picture test.

The test consists of five schematic instrument dials projected on the screen having simple indicators or pointers moving continuously in an irregular manner. The five dials, designated A to E, are highly schematic analogues of certain indicators on an instrument panel (air speed, ball, horizon, fuel-ratio, turn). For each indicator, a region is clearly marked off in black where the reading is "wrong." The reading not so marked is "right." The task of the candidate is to scan the dials and to record on a worksheet which dials "go wrong," even momentarily, in each trial,

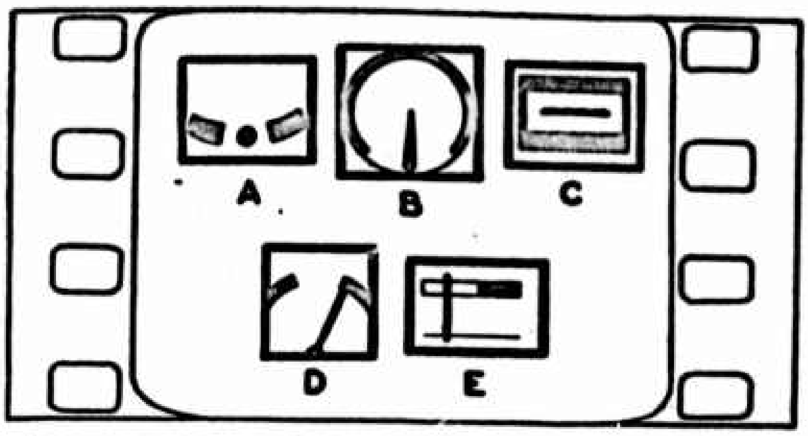


FIGURE 5.9.—Appearance of Screen During a Phase of the Flexibility of Attention Test. At the moment illustrated Dial D has gone wrong.

It is illustrated in figure 5.9. This information he later transcribes on a standard A-E answer sheet. The candidates are also instructed to pay particular attention to Dial B, and are told if they fail to record it when this dial goes wrong, or if it is recorded when it did not go wrong, that the error will count triple. The test was intended to arouse a considerable degree of "pressure," and the candidates are kept very busy taking note of the indicators that "go wrong." The test is subjectively difficult and demanding. In order to eliminate variance due to a presumably clerical type of skill in recording the dials neatly and exactly on the answer sheet, immediate recording is not required. Instead, the test is divided into 12-second "phases," and at the end of each phase the student is allowed a brief interval in which to write the letters of the appropriate indicators on a work sheet. The number of indicators which "go wrong" may vary from zero to five during a single phase. This number increases, in general, throughout the test. Several dials usually go wrong, and others threaten to do so, with mounting demand on the students' attention.

Details of Development. The test was photographed on 16 mm. film at Ft. Worth, while the Perceptual Research Unit was still active, by constructing a large (6 x 8 foot) indicator panel with hand-operated controls behind the panel which enabled the five indicators to be individually moved. The movements of each indicator were elaborately planned and specified second by second for each phase of the test, and five operators were rehearsed with a metronome in performing their tasks. The "go wrongs" were tried out and revised to give proper levels of difficulty, and finally photographed. Several forms of the test, including different methods of scoring, were tried out. The method of using work sheets was devised.

A study was made of the errors occurring when the candidates transcribe their responses from the work sheets to the answer

sheets. The study showed the number of errors occurring during transcribing to be small. It was concluded that this somewhat novel method was satisfactory. An invitation to the candidates to criticize the test got few responses, indicating the test and conditions of administration were reasonably satisfactory from their point of view. A study was also made on the effect upon transcription errors of using ruled versus unruled work sheets. No significant difference in number of transcription errors made on the different types of sheets was found.

In the instructions given the candidates preparatory to taking the test, they are told to pay especial attention to Dial B, as it is a critical dial and errors will be counted triple. However, since the correlation between the number of errors made on Dial B and the other dials was found to be high, it was decided that an actual weighting of this dial in scoring the test was not necessary.

The scoring of this test presented a situation unique in psychological testing, since an error could consist of either a response or failure to make a response, and a correct answer could either be a response or the proper omission of a response. Thus, a candidate could err by failing to record a dial going "wrong" or by recording a dial as having gone "wrong" when it actually had not. It was determined that the correct scoring formula was R–W, in which "Rights" and "Wrongs" are defined in terms of this double category. The items of the test consist of both seeing indicators go wrong (rightly or wrongly) and seeing them *not* go wrong (rightly or wrongly). There are therefore theoretically five items in each phase. A simplified scoring formula could be derived, however, from the above considerations in which "Rights" are indicators correctly marked as "going wrong" and "Wrongs" are indicators incorrectly marked as "going wrong." The formula is still R–W. The number of items in the test conceived in this way is fewer than five per phase.

Test Characteristics

1. Administration and Scoring:

Form	*Running time* *Minutes*	*No. of phases*	*Scoring Formula*
CP411A	10	16 (38 items)	R (Feb. 1943)
CP411B	13	26 (99 items)	R-W (Mar. 1943)
CP411C	16	20	R-W (Mar. 1943)
CP411D (Integration of attention	13	26 (76 items)	R-W (Mar. 1943)
CP411E	15	36 (137 items)	R-W (Mar. 1943)

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	r_e	*Type*	*Remarks*
A	2/43	Hq. TRC.	Local personnel	22	...	.73	Odd-even	
A	2/43	Hq. TRC.	Local personnel	22	...	.83	Odd-even	2d Admin.
A	2/43	Hq. TRC.	Local personnel	22	...	.53	Test-retest	Dial B weighted 3.
B	6/43	SAACC	Unclassified	277	...	.94	Hoyt	
E	12/43	SAAAB	Classified pilots 44-F	401	.50	.67	Odd-even	

3. Distribution Constants:

Form	Date	Place	Group	N	M	SD	Remarks
A.....	2/43	Hq. TRC	Local personnel......	22	37.7	8.3	Dial B weighted 2.
B.....	5/43	SAACC	Unclassified..........	902	88.5	8.56	
B.....	5/43	SAACC	Unclassified..........	852	90.8	9.59	Preceded by form C.
C.....	5/43	SAACC	Unclassified..........	933	54.78	9.26	
B.....	6/43	SAACC	Unclassified..........	505	90.14	8.39	
B.....	6/43	SAACC	Unclassified..........	528	94.10	8.94	Preceded by form C.
E.....	12/43	SAAAB	Classified pilots, 44-F	1,376	116.84	7.38	

4. Validity:

Form	Date	Place	Group	Type	Nt	P_p	M_p	M_f	SD_t	r_{bis}	[illegible]	r_{tet}
B....	6/43	SAACC	Unclassified...........	Bis	295	.808	94.55	93.83	8.7	.05	...	.25
B....	6/43	SAACC	Unclassified...........	Bis	219	.740	91.56	87.61	9.0	.26	...	.23
E....	12/43	SAAAB	Classified pilots, 44-F.	Bis	1,097	.906	117.60	115.55	7.4	.15	.20	.22

The criterion used was graduation-elimination from elementary pilot training.

5. Intercorrelations :

Correlated with	Date	Place	Group	N	r
Integration of attn. CP411D.......	6/43	SAACC	Unclassified..........	277	.27
Integration of attn. CP411D.......	12/43	SAAAB	Classified pilots 44-F.	1,376	.26

The intercorrelations obtained in the June 1943 study are shown in the following table. The upper figures in each row are for order BD (form B given first), N = 477; the lower figures are for order DB, N = 494. The cases were unclassified cadets, tested at San Antonio Aviation Cadet Center.

	1	2	3	4	5
1. Flexibility of Attention CP411B..........	..	.40	.35	.23	.25
		.31	.29	.23	.23
2. Integration of Attention CP411D........		..	.36	.25	.31
			.28	.23	.22
3. Bombardier Stanine......................			..	.71	.79
				.65	.79
4. Navigator Stanine......................				..	.44
					.39
5. Pilot Stanine............................					..

6. Practice Effect:

In the June 1943 administration at San Antonio Aviation Cadet Center half of the subjects were given Form B of the test first, followed immediately by Form D (Integration of Attention), while in the other half this procedure was reversed with Form D being given first. By comparing the mean for Form B when given first with the mean when given in second order, it was possible to determine the practice effect due to the prior administration of Form D. The results were as follows:

Form B (Flexibility of attention)

	N	M	SD	CR
Given before D...........................	477	90.14	8.39	7.12
Given after D............................	494	94.10	8.94	

Form D (Integration of attention)

	N	*M*	*SD*	*CR*
Given before B	494	39.58	9.78	18.85
Given after B	477	47.62	9.16	

In this case, taking Form D before taking Form B increased the scores 3.96 score points or .47 of a sigma. Similar results were obtained for Form D with practice on Form B increasing the score 8.09 score points or .83 of a sigma. Practice was thus shown to have a significant effect, with the scores of Form D showing an appreciably larger gain.

Integration of Attention Test CP415A. This test was designed to measure what might be called "integrative" attention. It was felt that the pilot must not only be able to distribute his attention over a complete field of events (as in Flexibility of Attention), but he must be able to treat this field as an interconnected whole. That is, he must see, and react to, patterns and combinations of events and relations between events, rather than single events. A particular set of events occurring at the same time (instrument readings, external cues, or both) requires an immediate response appropriate to that combination.

This test employs the same content as Flexibility of Attention, but requires the subject to look for and record *simultaneous combinations* of wrong dials. In addition to watching all the dials, he must perceive them as a whole in order to note the combinations which are formed, as illustrated in figure 5.10. Isolated wrong dials do not count.

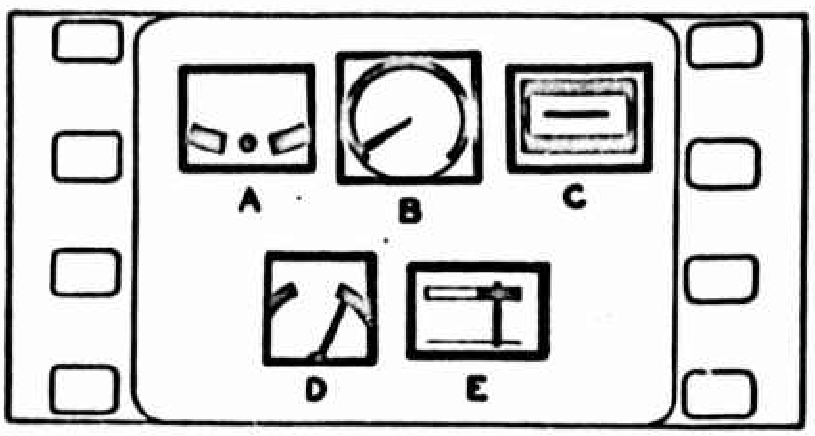


FIGURE 5.10.—Appearance of Screen During a Phase of the Integration of Attention Test. At the moment illustrated three dials have gone wrong simultaneously.

Details of Development. The first form of the test was made by using the film of Form B, Flexibility of Attention, which showed enough simultaneous combinations to use for a tryout. It was at first called Form D. The new task proved to be dissimilar to the earlier one and constituted a new test. It was therefore given a

new name, and a new film, specially designed to show combinations of indicators, was photographed with the panel described. The combinations become very complicated in later phases of the test. The test has the quality of appearing to require, while undergoing it, an ability which any alert person ought to possess and seems to arouse strong ego feelings. Many candidates experience a sense of inadequacy, whether or not they are performing well in a statistical sense.

The test was constructed and is scored in a manner similar to that already described for the Flexibility of Attention Test.

Test Characteristics

1. Administration and Scoring:

Form	*Running time*	*Number of phases*	*Scoring formula*
	Minutes		
CP411D...........	13.................	26 (76 items).................	R-W (June 1943).
CP415A...........	16.................	36 (104 items).................	R-W.

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	r_c	*Type*
D...........	6/43	SAACC	Unclassified.................	277	.88	.	Hoyt.
A...........	12/43	SAAAB	Classified pilots 44-F.........	400	.71	.83	Odd even.

3. Distribution Constants:

Form	*Date*	*Place*	*Group*	*N*	*M*	*SD*	*Remarks*
D.....	6/43	SAACC	Unclassified.........	365	39.54	9.65	
D.....	6/43	SAACC	Unclassified..........	296	48.34	8.29	Preceded by CP411B.
A.....	11/43	SAAAB	Classified pilots 44-F.	1,097	74.88	11.25	

4. Validity:

Form	*Date*	*Place*	*Group*	*Type*	N_t	P_g	M_g	M_e	SD_t	r_{bis}	σr_{bis}	r_{tet}	*Remarks*
D....	6/43	SAACC	Unclassified	r_{bis}	365	.808	39.75	38.64	9.65	.07	..	.31	
D....	6/43	SAACC	...do......	r_{bis}	296	.740	48.89	46.77	8.28	.15	..	.22	Preceded by CP411B.
A....	11/43	SAAAB	Classified pilots....	r_{bis} 44-F	1,097	.906	75.07	73.00	11.25	.09	.13	.18	

The criterion used was graduation-elimination from elementary pilot training.

5. Intercorrelations:

Correlated with	*Date*	*Place*	*Group*	*N*	*r*
Flexibility of attention CP411B..	12/43	SAAAB	Classified pilots, 44-F..	1,376	.26
Pilot stanine.....................	12/43	SAAAB	Classified pilots, 44-F..	1,097	.18

Other intercorrelations are given in preceding sections, Flexibility of Attention CP411E. It may be noted that the correlations of this test with Flexibility of Attention is only .26 despite the fact that the two tests have essentially the same content.

Sequential Perception

Perception of a continuous flow of events, to which an adjustment or response is made only by apprehending the sequence as such, is obviously not only important for complex locomotor behavior but also is a function uniquely adaptable to motion picture presentation. A simple method of presenting sequential stimuli

which have to be apprehended as a whole was suggested for use in the AAF testing program by Dr. Victor Lowenfeld of Hampton Institute. It consisted of moving a geometrical shape behind a slot in a cardboard screen and testing for correct perception of the shape. It was adapted for a motion picture test and another test involving sequential perception of a pattern was also devised.

Successive Perception Test I CP509C-I. This is a test of the ability to integrate successive partial impressions into a single visual scheme or pattern. Aircrew members, particularly the pilot, are often called upon to make appropriate responses to visually perceived situations of considerable extent and complexity, which

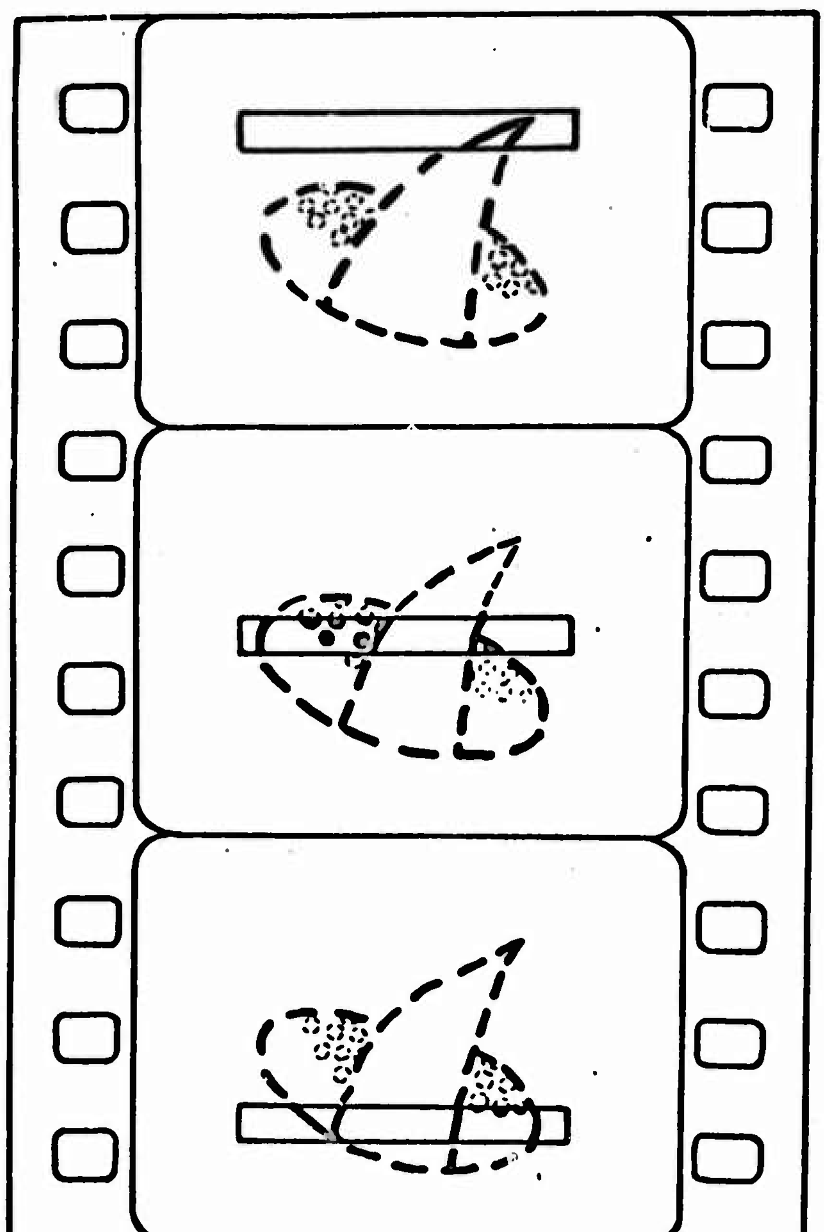


FIGURE 5.11.—Stages in a sample item of Successive Perception Test I. The dotted lines are not seen in an actual test item.

are not subject to view in their entirety at any one time. In other words the pilot or other aircrew member must be able to synthesize a clear total impression from brief successive glimpses of various parts of a situation. It is believed that this test might measure a "visualization" ability, which factor analyses have indicated to be important in predicting pilot success.

The test is composed of 45 motion picture items in which a slot in an opaque screen moves over a black pattern on a gray background, exposing it successively from top to bottom. These stages are illustrated in figure 5.11. The pattern is different and varies in complexity from item to item. After the pattern has been presented in this manner, it is shown again on the screen in its entirety along with four other similar patterns which act as confusion figures. The five patterns are labeled A, B, C, D, and E, as in figure 5.12. Candidates indicate recognition of the one

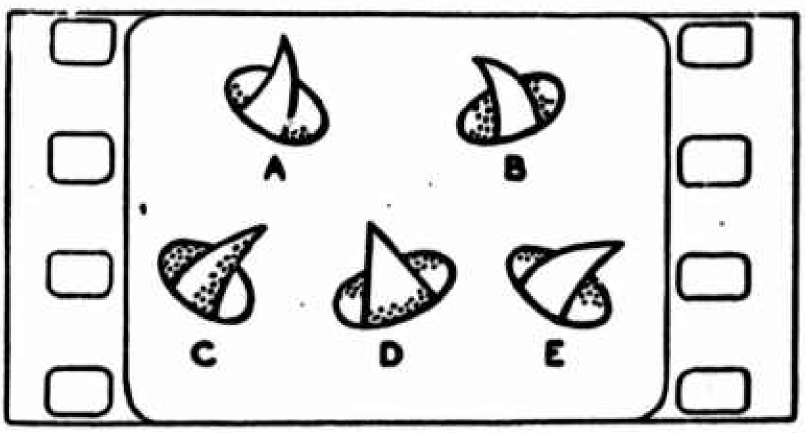


FIGURE 5.12.—The Five Alternative Responses Corresponding to the Sample Item of Figure 5.11.

which was presented successively behind the moving slot by marking in the appropriate space on a standard answer blank.

Details of Development. This test was originally photographed on 16 mm. film in the summer of 1943 at the Perceptual Research Unit in Fort Worth. Production was of a preliminary nature, and it was not intended that prints made therefrom should be a final form of the test. During the first months of 1944 the test was completely revised at the Psychological Test Film Unit all items were redrawn, and many new items were added. The test was photographed by animation on 35 mm. film at the First Motion Picture Unit in Culver City. Two experimental administrations were conducted while the test was in this form (Form B). Item analyses indicated that its reliability could probably be improved by elimination of a number of items. The final 16 mm. form of the test contains 38 items, including three practice items.

While Form B of this test existed in workprint form, an experi-

ment was undertaken to determine the effect of length of response-interval on reliability of total score. Two experimental forms of the test were made for this purpose: one in which the response alternatives remained on the screen for five seconds, and another in which they remained for 8 seconds. Reliabilities for the two forms were found to be .555 and .561. The difference is not significant.

Test Characteristics

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula.*
	Minutes		
CP509A-I	12	22	R.
CP509B-I	18	45	R.
CP509C-I	14	38	R.

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	r_e	*Type*	*Remarks*
B	8/44	SAAAB	Unclassified preflight.	416	.34	.50	Odd-even	
B	8/44	SAAAB	do.	415	.41	.59	...do...	Best 30 items.
B	8/44	SAAAB	do.	416	.56		Rulon	5-sec. response period.
B	8/44	SAAAB	do.	448	.55		...do...	8-sec. response period.

3. Distribution Constants:

Form	*Date*	*Place*	*Group*	*N*	*M*	*SD*
B	8/44	SAAAB	Unclassified preflight...	416	21.63	5.52

4. Validity:

No data available for inclusion in this report.

5. Intercorrelations:

Form	*Correlated with*	*Date*	*Place*	*Group.*	*N*	*r*
B	Successive Perception II CP509B-II	9/44	SAAAB	Unclassified preflight.	75	.18
B	Pilot stanine	9/44	SAAAB	do.	411	.04
B	Bombardier stanine	9/44	SAAAB	do.	416	.22
B	Navigator stanine	9/44	SAAAB	do.	414	.27

Successive Perception Test II CP509C-II (Moving Spot Test). This is another test intended to measure the ability to form an integrated total impression of a visual experience which has been

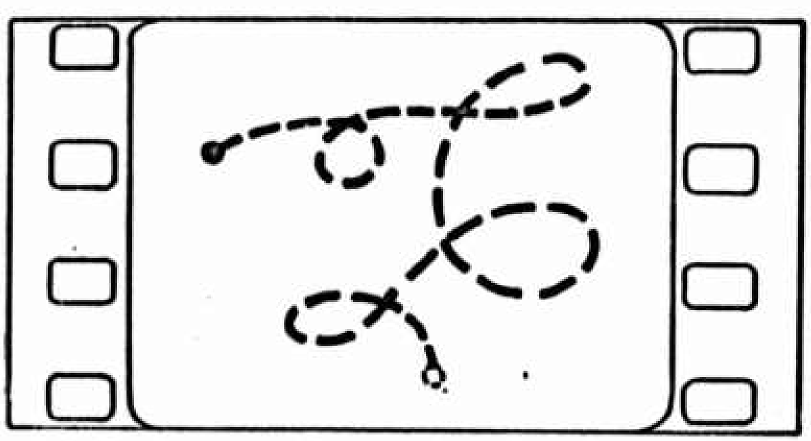


FIGURE 5.13.—Sample of Pattern Traced by Spot in Successive Perception Test II.

perceived in successive stages or parts. It was originally planned as an alternative to Successive Perception I but had to be regarded as a separate test of an independent function when it proved to correlate with that test only to the extent of .13.

The test is composed of 53 items (including three practice items) in which a black spot on a uniform light gray background moves over a complex path, thereby tracing an imaginary pattern, during an interval of ten seconds. A sample pattern is shown in figure 5.13. Immediately following the presentation of an item there are presented simultaneously on the screen five complex

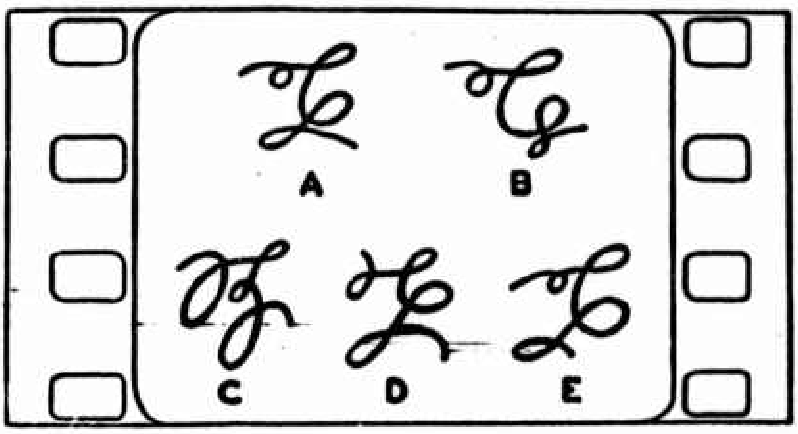


FIGURE 5.14.—The Five Alternative Responses Corresponding to the Sample Item of Figure 5.13.

patterns, one of which is identical with the path described by the moving spot, as in figure 5.14. The patterns are of three types:

a. Rectilinear (Part I in test; items 4–23).
b. Curvilinear (Part II in test; items 24–38).
c. Linear (Part III in test; items 39–53).

The patterns are labeled A, B, C, D, and E, and the subject indicates which pattern he believes to be the correct one by marking in the appropriate space on a standard answer sheet.

Details of Development. Form A of this test was composed of ten items on 16 mm. film, and was used only to make preliminary determinations of constructional problems. Form B was constructed in 35 mm. form by the First Motion Picture Unit, Culver City, California, and was administered to a group of unclassified Preflight candidates at Santa Ana Army Air Base in January 1944. Correlations were obtained between the three parts of the test, and between the whole test and Successive Perception Test I. These are reported in subsequent paragraphs. An item analysis was performed, and the fifty items having the highest phi coefficients with total test score were selected to be included in the final form of the test (Form C).

Test Characteristics

1. Administration and Scoring:

Form	Running time (Minutes)	No. trials	Scoring formula
CP509A-II	8	10	R.
CP509B-II	35	90	R.
CP509C-II	24	53	R.

2. Reliability:

Form	Date	Place	Group	N	r	r_t	Type	Remarks
B	6/44	SAAAB	Unclassified preflight	381	.52	.68	Odd-even	
B	6/44	SAAAB	do.	381	.34	.50	do.	Part I.[1]
B	6/44	SAAAB	do.	381	.37	.54	do.	Part II.[1]
B	6/44	SAAAB	do.	381	.32	.48	do.	Part III.[1]

3. Distribution Constants:

Form	Date	Place	Group	N	M	SD	Remarks
B	6/44	SAAAB	Unclassified preflight	381	47.72	8.38	
B	6/44	SAAAB	do.	381	16.12	3.52	Part I.[1]
B	6/44	SAAAB	do.	381	18.26	3.60	Part II.[1]
B	6/44	SAAAB	do.	381	13.44	3.70	Part III.[1]

[1]Part I was composed of curvilinear items; Part II of rectilinear items; and Part III of linear items.

4. Validity:

No data are available for inclusion in this report.

5. Intercorrelations:

Correlated with	Date	Place	Group	N	r
Successive perception I CP509B-I	9/44	SAAAB	Unclassified preflight	75	.13
Pilot stanine	6/44	SAAAB	do.	375	.06
Bombardier stanine	6/44	SAAAB	do.	376	.38
Navigator stanine	6/44	SAAAB	do.	375	.44
Part I vs. Part II[1]	6/44	SAAAB	do.	381	.40
Part II vs. Part III[1]	6/44	SAAAB	do.	381	.41
Part I vs. Part III[1]	6/44	SAAAB	do.	381	.32

[1]Part I was composed of curvilinear items; Part II, rectilinear; and Part III, linear.

Quickness of Perception

Speeded tests of form perception in printed form and using the method of recognition had proved to be predictive of pilot aptitude. It was believed possible that a purer measure of perceptual quickness could be obtained by a speeded presentation of each single pattern of such a test, and requiring actual reproduction of the pattern.

Plane Formation Test CP805C. This test was designed to measure the ability to apprehend a visual pattern within a brief exposure period and reproduce it accurately. Various considerations indicated that quickness of perception was an important factor in pilot success.

In order to isolate the factor of perceptual speed it was deemed best to use the tachistoscopic method and a motion picture was therefore prepared in which the number of frames of film determined the length of time for which the material was exposed. Material employed consisted of a grid of 25 squares upon which appeared 5 small plane silhouettes, as illustrated in figure 5.15.

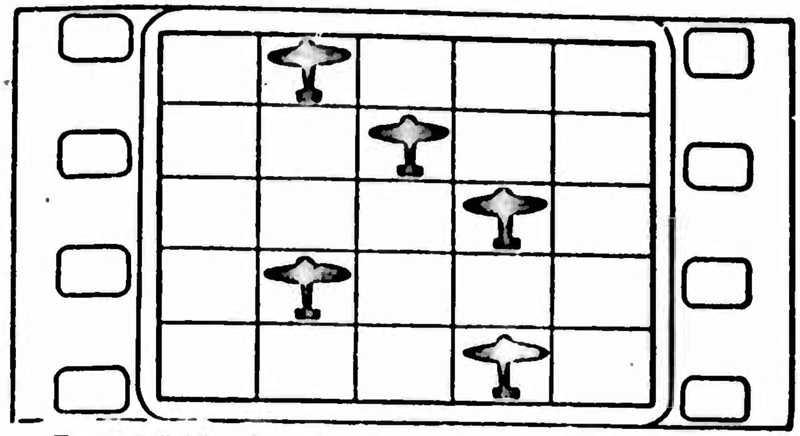


Figure 5.15.—Sample Item of Plane Formation Test.

The planes made a pattern or "formation." The task of the subject was to observe the screen and, immediately after the exposure, to fill in spaces on the answer sheet corresponding to the sections on the grid which included planes, i. e., to reproduce the pattern made by the five planes. An example is shown in figure 5.16.

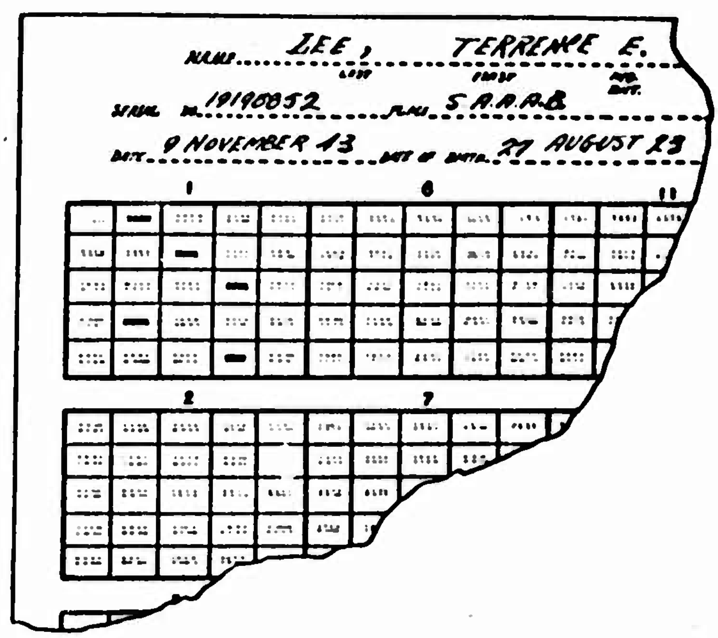


Figure 5.16.—Illustration of Answer Sheet Showing Correct Responses to the Sample Item Shown in Figure 5.15.

Details of Development. It was determined that the correctness of the reproduced patterns, when scored as patterns, was almost identical with the correctness of the reproductions when scored mechanically in terms of the number of grid spaces correctly filled in. If a subject got the pattern right he also got its position in the grid correctly. This made it possible to machine-score the answer sheets using the formula of "Rights" and counting each element of each pattern as one item. The test is a measure of

ability to reproduce a spot pattern, scored in terms of the correct location of each spot.

A preliminary form (Form A) was shot on 16 mm. film and tried out. Fifty patterns (150 items) were then photographed at Wright Field and administered to 400 aviation students. This was Form B. Item analysis was carried out, together with revision of the instructions and procedure, and the final test, Form C, on 16 mm. sound film, was produced using the best 30 patterns.

Test Characteristics

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula*
	Minutes		
CP805A................	18	38 (190 items).........	R (Feb. 1943).
CP805B.................	23	50 (250 items).........	R.
CP805C.................	15	30 (150 items).........	R.

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	*Type*
B.............	2/43	SAACC	Unclassified 43-K...	407	.82	Hoyt.

3. Distribution Constants:

Form	*Date*	*Place*	*Group*	*N*	*M*	*SD*	*Remarks*
B...........	2/43	SAACC	Unclassified 43-K...	407	107.1	13.0	Trials 1-30.
B...........	2/43	SAACC	Unclassified 43-K...	407	70.5	9.5	Trials 31-50.
B...........	3/43	SAAAB	Unclassified.......	460	173.89	21.86	Complete test.

4. Validity:

Form	*Date*	*Place*	*Group*	*Type*	N_t	P_g	M_g	M_e	SD_t	r_{bis}	r_{tet}
B ..	2/43	SAACC	Unclassified 43-K.....	r_{bis}	250	.640	180.09	172.94	20.04	.22	.34
C ..	6/43	SAAAB	Unclassified..........	r_{bis}	956	.785	112.45	109.77	13.73	.12	.26

The criterion used was graduation-elimination from elementary pilot training.

5. Intercorrelations:

Correlated with	*Date*	*Place*	*Group*	*N*	*r*
Estimation of velocity CP205B-I............	3/43	SAAAB	Unclassified.......	392	.05
Identification of velocity CP205B II.........	3/43	SAAAB	..... do.	392	.08
Estimation of relative velocity CP205B-III..	3/43	SAAAB	..... do.	392	.10

Comprehension of Visual and Vocal Instruction

Motion Picture Comprehension Test C1625A. This test was designed to measure the ability of aircrew candidates to comprehend and remember material which is presented in motion picture form with visual demonstration and diagrams accompanied by explanatory narration. Much of the instruction in flying training is given orally or by means of training films. A relatively small proportion of flying training is obtained by independent reading outside the classroom, especially in the advanced stages of flying training. For these reasons it seems possible that a test of comprehension presented by means of motion pictures, in which the material is presented both orally and visually, might yield a better

prediction of success in flying training than does the Reading Comprehension Test. There are some important differences between the Reading Comprehension Test and a Motion Picture Comprehension Test. In reading comprehension, skill in reading is an important factor in performance. This is not true in motion picture comprehension, where the verbal material is given orally. Second, in motion picture comprehension the subjects must answer the questions on the basis of one presentation of the material. They cannot refer back to the explanatory passages as they can in a reading comprehension test. Third, in a motion picture comprehension test the ideas and concepts are not presented merely as verbal symbols. It has been demonstrated that comprehension of ideas presented solely in words has low pilot validity. It is probable that some individuals who fail to grasp concepts explained verbally can adequately understand them when they are presented with animated diagrams, pictures, and illustrations, and that the two kinds of comprehensional ability are not highly correlated. The ability to comprehend relationships presented in this less abstract way may be important in pilot training. The test was originally suggested by Lt. A. F. Jenness of Medical and Psychological Examining Unit No. 7.

The test is made up of six technical sections or excerpts taken from training films, ranging in length from two to seven minutes. A passage is presented to the subjects, immediately followed by a series of questions on the material, similar to those used in the Reading Comprehension Test. The test has an overall length of about thirty minutes. Multiple choice questions for each of these excerpts were written, edited, and tried out in preliminary fashion, after which the satisfactory questions were selected. The six passages of the test are as follows:

Code No. of film		*Running time*	*No. of questions*
		Minutes	
T.F. 1-290	Celestial Nav.—Intro. and location of points on the celestial sphere, section I	3½	10
T.F. 1-290	Celestial Nav.—Intro. and Location of points on the celestial sphere, section II	3	7
T.F. 1-544	Interpretation of bearings, lines of position, and fixes, section I	3	7
A.F. 212B	Setting-up operations—C-1 Autopilot	2	6
T.F. 1-497	Methods of high level bombing, section I	4½	8
T.F. 1871	Norden Bombsight principles	7	8

The test exists in the form of a script which includes a 90-second passage on Night Bombing and three questions on this material for practice. The six passages above, specified in terms of the commentary accompanying the films, are written out and are accompanied by the questions on the vocal-visual material presented. The test can be administered in the form of a spliced set of the

excerpts cut from 16 mm. prints of the 6 training films in question, the questions being presented in the form of a booklet.

A version of the above test, the "Position Firing" Comprehension Test, utilizing a single 15-minute training film instead of excerpts from different training films was tried out in connection with another research project on the evaluation of motion picture instructional techniques. The film involved is an instructional picture for aerial gunners explaining the system of position firing (TF 1–3366). A 25-item examination on the complex spatial ideas, rules, and procedures to be learned by the flexible gunners in firing at attacking fighter planes was constructed and revised. The test measures comprehension of the film, as well as being a proficiency measure. Its characteristics are given here as probably being representative of the more general test described above.

Test Characteristics

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula*
Preliminary (Position Firing)..............	35 minutes	25	$R - \frac{W}{4}$

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	r_e	*Type*
Preliminary...	6/44	SAAAB	Unclassified Preflight....	332	.46	.63	Odd-even

3. Distribution Constants:

Group	*Date*	*Place*	*Group*	*N*	*M*	*SD*
Preliminary......	10/44	SAAAB	Unclassified preflight...	328	18.41	3.54

4. Validity:

   No data were obtainable for inclusion here.

5. Intercorrelations (Preliminary Form, "Position Firing")

Correlated with	*Date*	*Place*	*Group*	*N*	*r*
Bombardier stanine................	10/44	SAAAB	Unclassified preflight..	328	.28
Navigator stanine..................	10/44	SAAAB	....do.................	327	.31
Pilot stanine........................	10/44	SAAAB	....do.................	328	.18
Reading comprehension (Booklet 32, Part I)............	10/44	SAAAB	....do.................	328	.47

VALIDITY

No general statement can be made about the predictive validity of motion picture aptitude tests for success or failure in flying training. For a number of tests, the data necessary for computing validities could not be obtained before the termination of large-scale pilot training. This was due in part to the fact that the Psychological Test Film Unit did not itself administer, score, and keep records of tests given to aviation students on its own responsibility. It therefore depended on the cooperation of other units, having their own tests to administer, for the testing of very large

samples of the sort required for validity studies. Such cooperation was obtainable, but it necessarily involved administrative formalities and delays.

The first six tests completed (the three tests for estimation of visual speed, the Plane Formation Test, and the Flexibility and Integration of Attention Tests) were each given to one or more samples for validation against graduation-elimination from elementary pilot training. The biserial coefficients of correlation do not always agree from one sample to another. The highest coefficients obtained are, however, only moderate. They have been reported separately for each test.

The next five tests produced (The Minimal Movement and Drift Direction Tests, the two tests for Successive Perception, and in particular the carefully thought out Landing Judgment Test) were not given to large samples in time for the pass-fail data to mature. The giving of these tests on a large scale was delayed in the first place for some months because of the critical shortage of film with which to make extra prints in the summer of 1944, and later by administrative delay in the determination of validation policy. Efforts to validate some of these and other tests against specific criteria which the test was designed to predict, such as landing grades in primary pilot training, proficiency scores in synchronizing a bombsight-trainer, and instrument flying grades, did not meet with encouragement.

The last of the tests described (the Distance Estimation Test, the two orientation tests, and the Motion Picture Comprehension Test) were completed too late for validation.

Nine of the above motion picture tests were finally administered in a comprehensive battery of experimental aptitude tests at Medical and Psychological Examining Unit No. 6, Keesler Field, Mississippi, during the summer of 1945, for the purpose of studying their intercorrelations with other tests, for various other statistical studies, and for further study of the effect of seating position on scores. The best information about the characteristics of the above motion picture tests would be derived from the data of this comprehensive study, but they are not available for inclusion in this report.

From such evidence as is available, both the intercorrelations of motion picture tests, and their correlations with other tests, seem in general to be low. Some of these data are reported separately for each test. The low correlations with other aptitude tests are consistent with the theory that motion pictures are capable of testing functions not amenable to other forms of testing. The generally low intercorrelations between motion picture tests themselves, even when they were designed to measure theoretically similar functions such as speed estimation or sequential percep-

tion, indicate uniqueness (if the reliabilities are taken as correct). The fact that motion picture tests, as compared with printed tests, do not require verbal ability or the vocabulary factor implies that in this respect they have *less in common* than printed tests. The effect would be to lower the intercorrelations of motion picture tests. It is likely that there are types of human aptitude and ability, only touched upon by the tests described, which cannot be adequately measured by the relatively static problems and questions presented by ordinary test methods but which can be demanded by setting up tasks arising from the continuous flow of events portrayed on the motion picture screen.

CHAPTER SIX

Proficiency Tests*

Background of Motion Picture Proficiency Testing

In the latter part of 1943, at the time the Psychological Test Film Unit was being established, the efforts of the AAF Aviation Psychology Program began to turn from problems of selection exclusively toward problems of training. The first formal request for an experimental study of training methods by Headquarters AAF to the Training Command was made in September 1943. It concerned the procedures being used for training students in the recognizing of hostile and friendly aircraft. It also suggested that the best criterion of proficiency would be motion pictures of the planes to be recognized rather than a series of still pictures such as were currently being used for examinations. Motion pictures showing the changing aspects of a plane in flight would approximate the situation in which aircraft have to be recognized in combat. The planning and execution of this study were assigned to the Psychological Test Film Unit.

The series of experiments performed in carrying out this project will be reported in the next chapter. The motion picture examinations used to measure the effectiveness of the training, however, will be described here. They consisted of a proficiency test at the Preflight School level which was used in the training experiments; a standard general examination for all levels, in two equivalent forms, which was derived from it; and a specialized examination separating the aircraft of the European and Pacific theaters, also in two forms, for gunners.

Research on measuring and training the visual skill of recognition absorbed much of the effort of the Film Unit during 1944. At about the same time, training research was being carried out on a large scale in pilot training, navigation, bombardiering, gunnery, and finally in the training of radar observers. The measuring of levels of proficiency was a first necessity in these fields. In the case of several specialties, it appeared that the motion picture medium could be used to construct examinations which were more similar to the actual performances for which students were being

*The tests to be described in this chapter were principally the work of R. M. Gagne in collaboration with the editor. Two of the tests were joint projects with other aviation psychologists.

trained than could the conventional written examinations. The motion picture would, at the same time, present a standardized situation capable of being given to a whole group at once, in contrast to individual performance checks.

Research was pursued, therefore, on a number of motion picture proficiency tests in collaboration with the Psychological Research Projects for the specialties concerned. Two such tests were developed, one on practical aerial navigation, and another on the location and identification of bombing targets from the air. Preliminary plans for a proficiency test in radar navigation and radar bombing were never carried into effect.

In these proficiency tests, the effort was made to get away from the purely academic type of examination which puts a premium on verbal memory and to test the performance of the student in a situation having the sequence, the tempo, and the continuous change of the real situation with which he will have to deal. Academic instruction is commonly criticized for giving students the theory of a discipline but taking no interest in how it can be put into practice. One method of bridging the gap is the use of motion picture group testing. The five proficiency tests will be described in the order in which they were constructed.

Recognition Testing

The Aircraft Recognition Proficiency Test (Preflight Level)

A preliminary test was constructed in the following way. The 35-mm. film library of the AAF Motion Picture Unit was searched for shots of different airplanes in flight. A large accumulation of such shots was available for use in the series of training films being produced. Shots were chosen which showed the plane in actual flight against a natural background of sky, clouds, land, or distant horizon. Some were films taken from the ground and others by aerial photography. Shots were selected, whenever possible, showing the plane in a continuously changing aspect, i. e., in transitions between head-on, passing, and plan views.

From these shots, by inspection in a movieola, items were cut at the most advantageous part of the flight, from 2 to 5 seconds in length. These items were spliced together, with 6-second intervals of blank film between them to permit the recording of answers by the students. This preliminary form of the test included 31 American and British aircraft, both Army and Navy, which were at that time (December 1943) being taught in AAF Preflight Schools.

There were 100 items in the preliminary test, each plane being represented by three or occasionally four items. Of these, one was relatively easy to identify, another of medium difficulty, and the third was difficult. Some represented head-on views, which

are the most difficult to recognize, in general, and some were distant views. The planes were distributed in a random order throughout the test and the order of items ran from easy to difficult.[1] The test required 20 minutes to run off, with oral instructions given by the test administrator. It was used with a mimeographed answer sheet having spaces for recording 100 answers. Answers had to be written by the students (P-38, Spitfire, etc.) since the *ability to name the plane promptly* was considered essential in recognition training. All aircraft recognition examinations therefore had to be hand-scored. The method of response by selection from five alternatives, which would have permitted the use of machine-scorable answer sheets was tried but abandoned when it met with the opposition of recognition training authorities.

The preliminary form of the test was then revised by selecting 64 additional items representing enemy planes, both German and Japanese. The total of 164 views was tried out with groups of preflight students and recognition instructors at Santa Ana Army Air Base, and an item difficulty analysis was made. Further selection of items for proper type and level of difficulty reduced them to 100. The range of difficulty was from 20 percent to 95 percent correct—the items being distributed fairly evenly at all levels rather than concentrated at any one level. The planes finally included in the test were 46 percent American, 24 percent German, 19 percent British, and 11 percent Japanese. The interval for recording answers was increased to seven seconds. Instructions, by titles and sound track, were incorporated into the film and a number of 16 mm. sound prints of the test were produced. They were used as measures of achievement in experiments on different methods of recognition training. The test was entitled the Aircraft Recognition Proficiency Test (Preflight Level).

Test Characteristics

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula*
	Minutes		
Preliminary	20	100	Rights.
ARPT (Preflight level)	24	100	Do.

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	*rc*	*Type*
Preliminary	1/44	SAAAB	Pilot preflight	81	.90		Test-retest (3 days).
ARPT	3/44	SAAAB	do.	420	.81	.89	Odd-even.

3. Distribution Constants:

Form	*Date*	*Place*	*Group*	*N*	*M*	*SD*
Preliminary	12/43	SAAAB	Preflight 44-G	389	59.0	15.0
ARPT (preflight)	2/44	SAAAB	Preflight	250	59.10	11.59
ARPT (preflight)	3/44	SAAAB	Preflight	352	62.15	11.92
ARPT (preflight)	4/44	SAAAB	Preflight	1011	56.07	12.26
ARPT (preflight)	6/44	Primary	(6 schools)	685	53.30	13.70

[1] In the selection of shots for this and the following tests and in other ways assistance was given by Major G. Emmanuel, aircraft recognition expert of the British Army Air Staff, Washington, and by Lt. (later Major) R. Yeldham, of the AAF Training Aids Division.

703313—47—8

4. **Intercorrelations:**

Correlated with—	*Date*	*Place*	*Group*	*N*	*r*
Final slide exam.	3/44	SAAAB	Preflight P, 44-I	387	.55
Final slide exam.	4/44	SAAAB	Preflight P, 44-J	502	.61
Airplane photo test	2/44	SAAAB	Preflight B & N	234	.74
Pilot stanine	8/44	SAAAB	Class. Pilots	672	.25
Bombardier stanine	8/44	SAAAB	Class. Pilots	672	—.08
Navigator stanine	8/44	SAAAB	Class. Pilots	672	.06

Results

The Aircraft Recognition Proficiency Test, although designed primarily to serve as a criterion for experimental studies, turned out to be in many ways a better examination for the 30-hour, 8-week course in recognition in the Preflight School than the examination actually in use. The latter was, by Training Command regulation, a 40-item series of photographic slides which were flashed on the screen showing views of the airplanes learned in the course. Frequently they were the same slides that had been used in training. Seventy percent of the slides had to be identified correctly to obtain a passing grade.

Because of this requirement that 70 percent was the passing grade, instructors obviously could not first construct a good examination and then set the grades appropriately by an examination of the distribution of scores. Instead they had to juggle the selection of slides for the examination in such a way that all the students who deserved to pass could identify 70 percent or more of the slides. The consequence was that final examinations in aircraft recognition had a restricted range of scores and an unnecessarily low reliability.

At the request of the Recognition Branch of AAF Training Aids Division, a comparison was made of three types of examinations for aircraft recognition courses, the slide examination prescribed by Training Command Headquarters, the motion picture test, and a printed test consisting of 100 photographs of the planes taught which had been published experimentally by the Training Aids Division. Six sections (classroom groups) of aviation cadets at Santa Ana Army Air Base, totaling 250, were given all three tests at the conclusion of their 30-hour course. The distributions, means and standard deviations of the scores are presented in table 6.1. For comparison the slide examination scores have been given as percentages. The restricted spread of the scores on the slide examination is evident. The discriminating power of the photographic test and of the motion picture test, on the other hand, is satisfactory.

The corrected odd-even reliability of the photographic test was .90, and that of the motion picture test was .89. Both of these are satisfactory, but the reliability of the slide examination, although it was not computed, had to be lower because of the restricted

TABLE 6.1.—*Distribution of scores for three tests of recognition*

Scores in percent	*Number of individuals having scores at each level*		
	Slide examination	*Airplane photo*	*ARPT (movie)*
100	34		
95–99	53		
90–94	56		
85–89	48	8	2
80–84	36	15	6
75–79	16	29	15
70–74	12	34	29
65–69	3	50	34
60–64	3	46	40
55–59		37	34
50–54		21	38
45–49		10	24
40–44		7	18
35–39		3	11
30–34		2	2
25–29			1
20–24			1
	Percent	*Percent*	*Percent*
Average	88.8	64.3	59.1
SD	8.7	10.7	11.9

range of scores and the smaller number of items. The correlations of the slide examination with the photo test and the motion picture test were .50 and .56 respectively. These correlations are low estimates of the true relationship between the tests because of the restricted range of scores of the slide examination; nevertheless they suggest that there is a difference between the type of proficiency measured by flashing pictures briefly on a screen and by the other two methods. The correlation between the motion picture test and the photo test was .74 which implies that these tests are also not measuring quite the same kind of proficiency. Since the motion picture test presents stimuli more similar to real planes in a real sky it could be argued, on *a priori* grounds at least, that the type of proficiency it measures should have a validity superior to that of the others.[3]

Application to Training. As a part of the efforts to improve recognition training originating at Headquarters, Army Air Forces in Washington, during 1943 and 1944, consideration was given to the use of standardized examinations at the various stages of training for all AAF schools. The tradition of the 70 percent passing mark proved to be highly resistant to change however, no less in the Army than in civilian educational circles. Neither standardized motion picture tests nor standardized printed photographic tests ever displaced the slide examinations as the final determiners of grades in recognition courses. On the initiative of AAF Headquarters, however, the motion picture test of aircraft recognition was used for surveying proficiency at all levels of training and in all schools in order to obtain information as to the effectiveness of the courses being given. For this purpose, another form of the test was constructed.

[3]Efforts were made at a later period to obtain some measure of proficiency in recognition as demonstrated in combat operations with which tests could be validated. No satisfactory measure was obtainable.

The Aircraft Recognition Proficiency Examination (Forms A and B).

A great deal of care had been taken to construct the Aircraft Recognition Proficiency Test with such a distribution of item difficulty as to make the test discriminating at different levels of proficiency. It therefore proved to be adequate as an examination above the preflight level for which it was designed. It was concluded that a single proficiency test (in equivalent forms) could be used at all levels of recognition training.

As a result of a conference in April 1944 with the officer in charge of recognition training at AAF Headquarters in Washington and with a representative of the AAF Training Aids Division, plans were made to revise the test for use as a standard universal measure of recognition proficiency. The new test was to be named the Aircraft Recognition Proficiency Examination. It was a 16 mm. sound film, like the previous form, full instructions being given by sub-titles and by the voice of an unseen test administrator. Three practice trials were given before the test proper began. The sub-titles made it possible to administer the test with a silent projector if necessary. Each item number appeared on the screen and was announced by the voice just preceding the shot of the aircraft to be identified.

The planes covered by the test, in one or more views, were as follows:

United States		*British*	*German*	*Japanese*
P38	F4U	Spitfire	Me109	Zeke.
P39	TBF	Hurricane	Me110	Hamp.
P40	C46	Typhoon	FW190	Val.
P47	C47	Mosquito	Ju88	Sally.
P51	C54	Beaufighter	He111	
A20	B17	Wellington	Ju87	
B25	B24	Lancaster		
B26	B34	Halifax		

A two-minute rest period is provided between items 50 and 51. The projector is switched off during this interval. The equivalent form B of the test was constructed by making a new order of the items in Form A and printing 41 of them in reverse; i. e., reversing the picture photographically from right to left. Thus, Form B contained exactly the same views as Form A, and scores on the two forms should be exactly comparable. The use of these two forms in testing different groups at the same school or on successive hours or days prevented an artificial increase in score as a result of specific "coaching" of one group by another. It also minimized an artificial increase in score, on retesting the same students, caused by the students remembering the item numbers of specific planes. The key for scoring the test was printed on the film itself, both at the beginning and at the end. These lists did not show when the film was projected.

Test Characteristics

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula*
ARPE (Forms A & B)......	24 minutes...............	100	Rights.

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	r	r_c	*Type*
ARPE	4/44	SAAAB	Pilot preflight..........	283	.73	.84	Odd-even.

3. Distribution Constants:

Form	*Date*	*Place*	*Group*	*N*	*M*	*SD*	*Remarks*
ARPE	4/44	SAAAB	Preflight P. 44-J......	1041	56.07	12.26	
ARPE	7/44	SAAAB	Preflight P. 45-B......	1856	57.00	11.78	
ARPE	8/44	WFTC	Primary...............	1501	56.60	13.75	11 schools.

4. Validity (as predictor of success in pilot training):

Date	*Place*	*Group*	*Type*	*Criteria*	N_p	N_e	M_p	M_e	SD_t	r_{bis}	σr_{bis}	r_{tet}
8/44	SAAAB	Preflight 44-J	r_{bis}	gr.-el.	601	71	57.25	53.24	11.41	.18	.26	.25

It may be seen that, whatever the validity of this test for its intended purpose, it has a somewhat surprising capacity to predict success or failure in primary flying. The above tryout was given to aviation cadets at the end of preflight training. The test was presumably measuring, over and above proficiency in aircraft recognition, interest in aviation and the ability to learn and remember visual forms.

Survey of Proficiency in Aircraft Recognition. A large number of prints of the test described were manufactured, 75 of Form A and 75 of Form B, and detailed instructions were written for a booklet to accompany each film on procedures for administration of the test. In accordance with a directive from Headquarters, Army Air Forces in Washington, a cross-sectional and longsectional survey of proficiency was conducted throughout the AAF Training Command, in October 1944, at all schools and at all stages of training. A sample of 19,000 students were examined, distributed in 90 different schools.

The test was administered by recognition instructors in the various schools to trainees who had completed the recognition training prescribed for each stage. In certain cases the amount of time which elapsed between the final class period and the testing period was great enough for the scores to be appreciably influenced by forgetting; these instances were therefore omitted from the results insofar as comparisons between stages of training were concerned.

The average scores and standard deviations obtained at each stage of training are shown in table 6.2. The relationships are shown graphically in figure 6.1, where the thickness of the cross-hatched horizontal bars represents the concentration of aircraft recognition training within each training period, and the

TABLE 6.2.—*Average score on aircraft recognition proficiency examination at each stage of training*

Date	Group	*N*	Average	SD
Nov. 1944	Preflight (2 schools)	1,442	57.43	12.37
Nov. 1944	Primary (5 schools)	1,336	67.43	11.76
Nov. 1944	Basic (9 schools)	2,570	69.77	11.39
Nov. 1944	Advanced (21 schools)	3,186	71.52	11.16
Nov. 1944	Transition (3 schools)	559	68.10	11.75
Nov. 1944	Navigator (3 schools)	1,588	56.16	12.92
Nov. 1944	Bombardier (7 schools)	1,329	61.87	13.85
Nov. 1944	Technical (6 schools)	2,401	55.70	17.74

heights of the solid bars represents mean proficiency scores for each type of school.

All adjacent differences shown in the table between successive stages of training and between different types of schools are statistically significant. In general, the results show that proficiency as measured by the Aircraft Recognition Proficiency Examination is closely related to the amount of aircraft recognition training which the trainees have received. For example, the average score in Technical Schools (55.70) where a 24-hour course is given, is lower than that in Preflight Schools (57.43) where 30 hours of aircraft recognition training are given. This result, however, may be due in part to superior aptitude of preflight students. Because of such potential differences in aptitude, the safest comparisons are those which run horizontally in figure 6.1 As more and more training is given throughout pilot training, average proficiency scores rise. The fact that the average drops a little in transition training supports the conclusion that *the amount of training per unit time*, rather than the amount itself, is of greater importance for the maintenance of recognition proficiency. It will be seen that proficiency measured at the end of transition training is somewhat lower than that in advanced training; this fall is associated with a drop in amount of recognition training from 6 hours to 2 or 3 hours, distributed within a 10-week period. Another comparison which may be made is the following: In advanced bombardier schools, where the concentration of aircraft recognition training during the 18-week course is a little less than that in primary schools (.5 hours compared with .6 hours per week) the average test score is comparable to, though lower than, that of primary schools. In contrast to this, recognition proficiency is not entirely maintained in advanced navigator schools, where the amount of instruction is approximately 7 hours spread over a period of 20 weeks (.3 hours per week). This comparison is valid only if the aptitude of bombardiers and navigators for recognition is equivalent, which seems reasonable. The aptitude of pilots may well be greater for the study of aircraft recognition.

In general the results indicate that a certain minimum concentration of training per unit of time is necessary if the level of proficiency once reached is to be maintained at the same level. If the amount of forgetting is not counteracted by a sufficient amount of

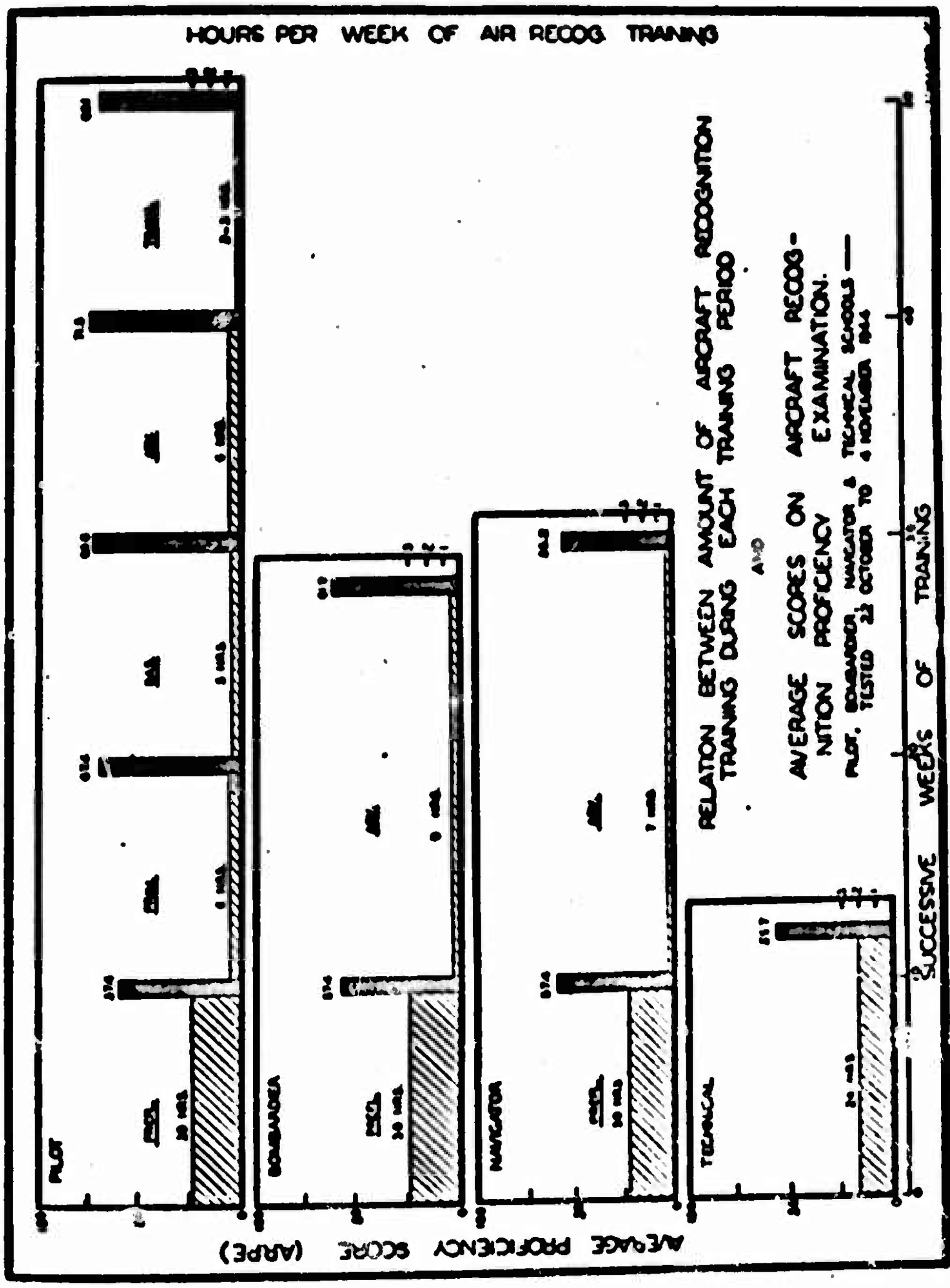

HOURS PER WEEK OF AIR RECOG TRAINING
RELATION BETWEEN AMOUNT OF AIRCRAFT RECOGNITION TRAINING DURING EACH TRAINING PERIOD
AND
AVERAGE SCORES ON AIRCRAFT RECOGNITION PROFICIENCY EXAMINATION.
PILOT, BOMBARDIER, NAVIGATOR & TECHNICAL SCHOOLS — TESTED 23 OCTOBER TO 4 NOVEMBER 1944
SUCCESSIVE WEEKS OF TRAINING
AVERAGE PROFICIENCY SCORE (ARPE)
PILOT
BOMBARDIER
NAVIGATOR
TECHNICAL

FIGURE 6.1.

relearning (or review) then a decrease in proficiency is the natural result.

The Aircraft Recognition Proficiency Examination for Flexible Gunners (Forms A and B).

Recognition training for aerial gunners was somewhat more specific than that for other members of the aircrew or that for fighter pilots. It was concentrated on the identification of pursuit planes, friendly or hostile, which would be likely to appear in the neighborhood of the bomber being defended. Moreover there was no need to learn the planes which would not be encountered in the theater of operations to which the gunner was assigned. A standard achievement test for gunners would therefore differ from the ARPE already described. A test, in two equivalent forms, was developed at the request of Headquarters, AAF, 30 prints were manufactured, and prints were distributed to the gunnery schools, with instructions for their use prepared by the Psychological Test Film Unit. The test was in other respects similar to those previously described.

The test was divided into two parts. Part I contained views of planes encountered in the European Theater; Part II had views of planes of the Asiatic-Pacific Theater. The following is a list of planes covered by the test:

Part I (European theater)			Part II (Asiatic-Pacific theater)			
United States	*British*	*German*	*United States*		*British*	*Japanese*
P38	Spitfire	Me109	P38	F4F	Spitfire	Zeke.
P39	Hurricane	Me110	P39	F6F	Hurricane	Hamp.
P40	Typhoon	Me210 (410)	P40	F4U	Typhoon	Oscar.
P47	Mosquito	Ju88 (188)	P47	TBF	Mosquito	Tony.
P51	Beaufighter	Do217	P51		Beaufighter	
		He111				

The test items were selected from an original set of 170 views. These were given as a test to several groups of aircrew trainees previous to their graduation from preflight training, and to a group of recognition instructors. Essentially the same criteria were used for selection of items as have been described for the Aircraft Recognition Proficiency Examination. Interviews with combat gunners, and a collection of opinions from officers of Psychological Research Unit No. 11 who had made job analyses of combat gunnery, led to the conclusion that the views of planes most similar to those seen in combat would conform to the following specifications: (a) a high percentage of "medium distant" and "distant" views; (b) approximately 10 percent plan views, 20 percent head-on views, and 70 percent passing views; including three-quarter and one-quarter views. These specifications were fulfilled to as great an extent as possible within the limits of the available material. In addition the views which depicted the greatest amount of movement and changing attitude were selected for the final test. In the case of German and Japanese planes particularly, some

views of model planes and some taken with a gun camera were used because they were the only ones available. Form B was constructed by rearranging the order of items within each half of the test, and reverse-printing a total of 39 items.

The use of this test as a proficiency measure was supervised by the Reasearch Division of the Central School for Flexible Gunnery rather than by the Film Unit. The results of its use were not reported to the Psychological Test Film Unit and cannot be summarized here. The data below were obtained with aviation cadets completing pilot preflight training.

Test Characteristics

1. Administration and Scoring:

Form	*Running time*	*No. trials*	*Scoring formula*
A, B	23 minutes	100	Rights.

2. Reliability:

Form	*Date*	*Place*	*Group*	*N*	*r*	*rc*	*Type*	*Remarks*
B	7/44	SAAAB	45-B	183	.78	.87	Odd-even	Test of 96 items.

3. Distribution Constants:

Form	*Date*	*Place*	*Group*	*N*	*Mean*	*SD*
B	7/44	SAAAB	45-B preflight	183	61.6	10.75

Other Proficiency Tests

Navigation Proficiency Test (Map Reading and Dead Reckoning)

This test was constructed to provide a standard measure of achievement in pilotage and dead-reckoning navigation. It presents, in motion picture form, a practice navigation mission which can be "flown" in a projection room by a group of navigator students. It has the advantage of retaining the natural tempo of a real navigation mission and the pace of the navigator's tasks when actually flying over unfamiliar terrain, together with the advantage of a standardized repeatable set of problems which may be given as a group test. It was developed in cooperation with the Psychological Research Project (Navigator), Ellington Field, Texas, and was intended primarily for the use of that project. It was administered to groups of navigation trainees at Ellington Field and the scoring procedures were developed there. The use of the test is described in the comprehensive report entitled *Psychological Research on Navigator Training*, chapter 5.

On the screen is shown a view of the moving terrain ahead from 12,000 feet, taken with the camera tilted downward about 45°. The view is that which would appear from an airplane during a typical two-leg navigation mission of 1½ hours' duration, covering about 300 miles. Superimposed on the picture at the bottom of the screen are a compass and air speed indicator, and numbers indicating

temperature and altimeter readings, each showing values *which correspond at all times to values appropriate for the actual mission flown by the plane which photographed the terrain.* Also superimposed is a driftmeter grid whose position is changed every 30 seconds, and underneath it the drift reading which corresponds to each setting. The examinee obtains correct information concerning heading, air speed, temperature, and altitude from the instrument readings shown. In the case of drift, however, the reading and setting are incorrect approximately half the time, and the examinee must take several readings in order to determine the actual drift during any portion of the mission. Two "double drifts" are flown, to make possible the computing of the wind vector, one during each leg of the flight. Examinees, seated in the testing room, are provided with stop watches, computers, and sectional aeronautical charts covering the route flown, and are required to keep a log of the mission. Scores obtained from this log provide the desired measures of navigation proficiency.

The film is composed of two 16 mm. silent reels, each about 1,650 feet long, and requires 90 minutes for administration. The test is run off on a standard sound projector at 24 frames per second (sound speed). Watches are stopped briefly during the rest period provided by the necessity of changing reels, and started again at the beginning of the second reel. Examinees are expected to identify checkpoints by comparing the terrain shown with the sectional maps provided. The strip of terrain appearing on the screen within which landmarks are identifiable is approximately 5 miles wide. At the bottom of the screen the terrain shown has a width of three and a half miles. The route flown (central California east of San Francisco) was carefully chosen as having checkpoints neither too difficult nor too easy for testing purposes. The kind of navigation required is precision dead reckoning aided by map reading.

The writing of exact specifications for this test, and the production of the film, presented a number of problems in the technical or scientific use of motion pictures. It required intimate collaboration between a psychologist familiar with navigation, a psychologist familiar with motion picture technique, the film cutter, and the producing studio, the AAF Motion Picture Unit. After the aerial photography had been accomplished with a camera plane, the instruments were separately photographed and superposed on the terrain film by optical printing. The instruments were synchronized with the terrain by a set of cues, at 30-second intervals, which correlated time and film-footage. Footage and time can be correlated exactly, since 35 mm. cameras and all projectors (at sound speed) are precision instruments. The instrument readings were specified in part from a record kept during the actual flight

of the camera plane, but chiefly by working backward from a navigation log of the flight performed.

Target Identification Test for Bombardiers

This test was designed to measure the proficiency of bombardiers, either those in training or those having completed it, in the identifying of bombing targets from the air at the time when the bomber is beginning its approach run and while the target is still a considerable distance away. The target for which the bombardier has been briefed must be located early in the run in order to allow time for the adjustment of the bombsight. Under certain conditions this task is very difficult; nevertheless the success of a bombing mission depends on it. Combat experiences indicated that there were wide individual differences among bombardiers in this skill, even after training, and that a test which would select those who were superior in target identification would be extremely valuable.

The test was originally proposed by the Psychological Research Project (Bombardier) and methods of reproducing the task of the bombardier on the motion picture screen were studied at length. Detailed plans of design and construction were worked out at a conference in February 1945 attended by representatives of both research organizations and of the Psychological Section in Training Command Headquarters. The plans were subsequently approved by Headquarters Army Air Forces.

The test consists of a number of approach runs at 20,000 feet on strategic target areas, the screen showing the bombardier's view ahead, and the task of the examinee is to locate as early as possible in the run the target on which he has been briefed. The run lasts about three minutes. Before each run the examinee is allowed one minute in which to study a marked map of the target region and a vertical aerial photograph of the specific target area. The course, the target, and all prominent reference points are indicated and are pointed out by voice recorded on the sound track.

Following this briefing period, the run begins. At first the target area is barely visible at the top of the screen, not far below the horizon, at a distance of some 20 miles. Identification is very difficult but general location may be noted if a reference point can be picked up. After 30 seconds of approach the progress of the run is stopped by the insertion of a "hold-frame" shot which remains on the screen for 4 seconds. Superposed on this still view of the terrain is a lettered grid, as shown in figure 6.2. The examinee now records on his answer sheet the space in which he believes the target to lie. The run continues for another 30-second period with another hold-frame shot which constitutes the next item. From three to six successive trials are run off in this manner, the loca-

tion of the target becoming progressively easier and more certain. When the target becomes identifiable by nearly all examinees, the run ceases. A new briefing period and a new target run on a different area begin.

The test was produced, after exploration of the possibilities of aerial photography, by photographing the large scale model of the Japanese islands which had been constructed at the AAF Motion Picture Unit for the production of training films which were used in the briefing of B-29 crews. The scale of this model (1 in. = 1,000 feet) was such that roads, factories, and terrain features were clearly visible and the scene approximated very closely to the reality when the camera action, kept to scale, simulated the course of a bomber over the terrain. The problem of obtaining "aerial" photographs and of describing landmarks exactly was simplified by employing this model.

The plans for this test included the photographing of runs on tactical targets taken from the air as well as its strategic industrial targets taken with the model. A section on the identification of ordinary checkpoints by map reading was also planned. Production of the test ceased, however, after the end of the war with Japan. Only a highly curtailed form of the test was completed, consisting of seven runs on the strategic targets which yielded 25 items. A few 16-mm. prints were manufactured, on silent film. The directions and the briefing periods were written in the form of a script to be read by the test administrator. Briefing is done with aerial photographs but without maps. The level of difficulty is judged to be satisfactory on the basis of informal tryouts. No data are available, however, from a formal administration of the test.

Administration and Scoring:

Form	*Running time*	*No. targets*	*No. items*	*No. scoring chances*	*Scoring formula*
A	25 minutes	7	25	50	Rights.

Conclusions

It would often be desirable in testing achievement or proficiency to present students with tasks which have the pace, movement, and continuity in time of the real tasks toward which their training is directed. This end can often be reached by using motion pictures. Five tests have been described which, in different ways, served this purpose. They suggest that examinations need not always be "academic" in character; they can be designed to measure knowledge or information as it is put into practice.

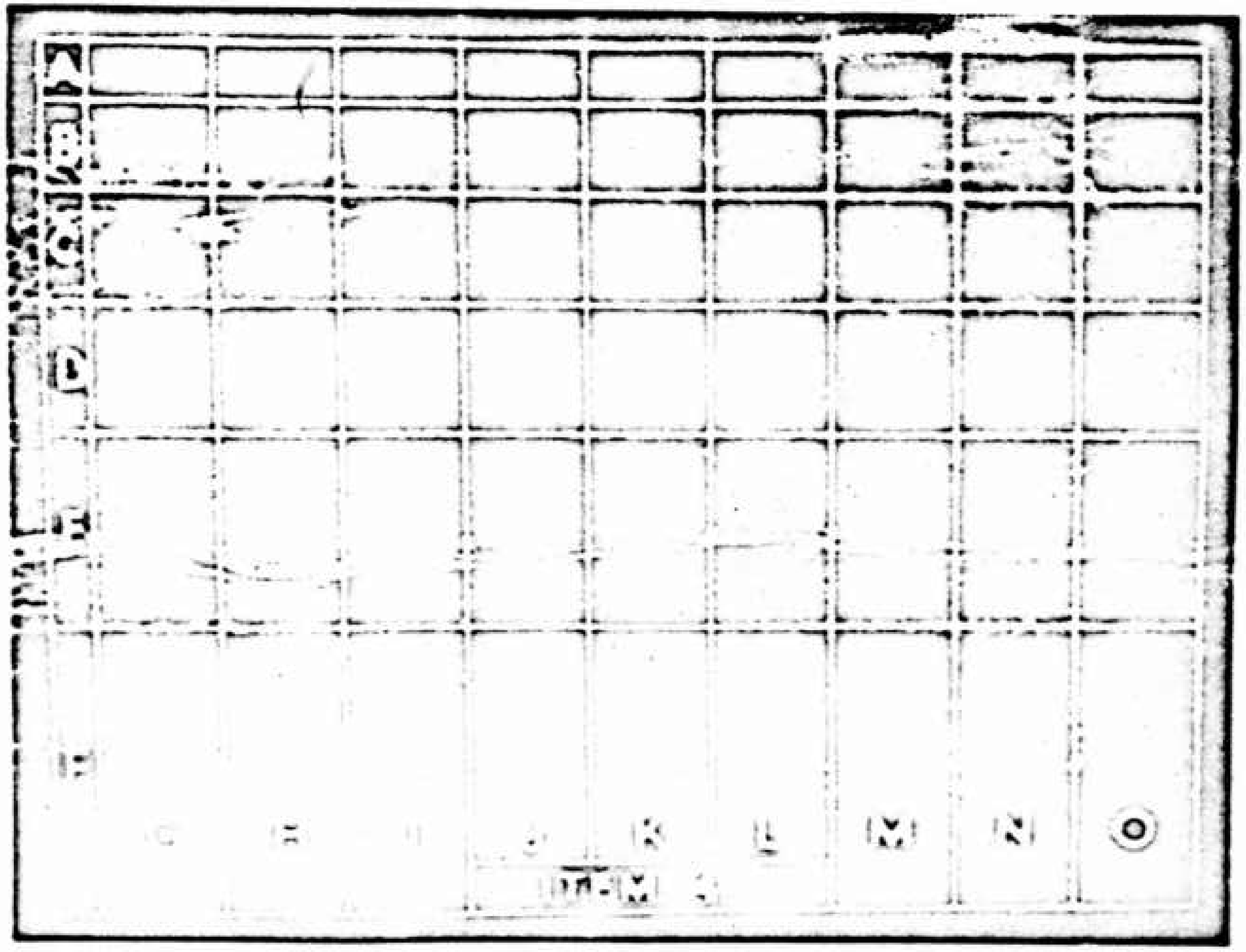


FIGURE 6.2.—Appearance of Screen During an Item of the Target Identification Test.

CHAPTER SEVEN

Research on the Recognition of Aircraft*

INTRODUCTION

Background of the Research Project

Training in the recognition of aircraft was a part of the curriculum in all AAF schools and was therefore given to substantially all personnel connected with aviation, both the combat crew and the ground crew. In addition, officers were attached to the units of the continental and overseas Air Forces who were responsible for the continued training in aircraft recognition of flying personnel engaged in operations. This type of training in the Eighth Air Force, for example, occupied at one time four hours per week. Since it was a universal subject, included in all specialties, it was carried out on a vast scale during the period when thousands of men per month were graduating from one school and going on to the next stage, and it occupied the time of a very large number of instructors.

In the latter part of 1943, stimulated by a number of dramatic instances of failure to recognize aircraft correctly in combat theaters, AAF Headquarters in Washington took steps to improve training in recognition. A monthly journal was published, having a wide distribution, and efforts were made to supplement and distribute training materials and training films through the AAF Training Aids Division. Moreover, the methods of instruction currently in use were called into question, especially the so-called "flash system," and the need was recognized of determining empirically the most effective methods of teaching the identification of aircraft.

A letter from the office of the Assistant Chief of Air Staff for Training to the Air Surgeon in August 1943, requested the services of aviation psychologists in determining the most effective methods. In accordance with this letter and a commendatory one which followed in February 1944, the Psychological Test Film

*The experimental research reported in this chapter was chiefly the work of R. M. Gagne. The experiments were designed by him in collaboration with the editor. The cooperation and assistance of many instructors and other officers in charge of training in aircraft recognition is acknowledged. The final draft of this chapter was written by the editor.

Unit was directed to initiate, and subsequently to continue, a program of experimental research on training in aircraft recognition. Some 10 or more studies were carried out, the results of which will be reported in this chapter. The principal measure of proficiency used for evaluation of training in the experimental and control groups, the motion picture test, has already been described in chapter 6.

Practices Employed in Teaching the Identification of Aircraft

History. Aircraft recognition as a subject for formal study was first systematically pursued in England in 1940, at the time when an air invasion was imminent. It aroused great popular interest and led to the training of a great body of civilian "spotters," in addition to being adopted as a subject for training in the Royal Air Force. As conceived by the British, the study of aircraft recognition included instruction about the nature and characteristics of different military planes over and above simple visual training in identifying them by shape and size. It thus reflected, in some degree, the general interest in aviation as such. Instructors in aircraft recognition needed to know a good deal about airplane design and about the most recent developments in aviation if they were to keep ahead of their students, whether military or civilian. The first school for training instructors in aircraft recognition was set up early in 1940 by the British Anti-Aircraft Command, and various types of training materials were developed, including manuals with photographic views of airplanes and the packs of self-study cards which later became familiar in this country. The prompt identification of an airplane as a friendly or enemy craft was a matter of vital importance to Britons in 1940, whether in or out of the military service.

When aviation training began its rapid expansion in the United States, more than a year later, aircraft recognition began to be taught as it had begun in England and went through a somewhat similar development. The fact to be noted is that the principal method of learning airplanes in these early days, so far as the visual aspects were concerned, was simply that of looking at pictures or silhouette-views (plan view, passing view, and head-on view) and attempting to memorize their appearance in connection with the name of the plane. The repetition of pictures and names in pairs with frequent attempts to identify, rightly or wrongly, and the correction of wrong responses—a method similar to the learning of paired associates by "prompting" in the psychological laboratory—was not at first utilized. The appearance of planes had to be committed to memory and it was easy to assume that this had to be done in somewhat the same way that a student would memorize a poem. Since all aircraft look more or less alike

and many look very much alike, they had to be memorized in considerable detail. A terminology, therefore, arose for the shape-characteristics of planes, borrowed in part from the designers, and there came into use an arbitrary order of memorizing these characteristics by wings, engines, fuselage, and tail. The letters W, E, F, and T represented this elementary attempt at systematization, and the method of learning came to be known as the "WEFT system." It was primarily an aid to memorization rather than a system of instruction. A student could study alone by this method without the help of an instructor or the use of special training procedures. It had the defect of over-emphasizing those aspects of the shapes of aircraft *which could be given names* ("swept back," "dihedral," "taper," etc.) to the neglect of other aspects not easily nameable. Furthermore, the aspects memorized were frequently not those by which the similar shapes could be discriminated from one another. The learning tended to be verbal in character—a list of characteristics which might or might not arouse an adequate visual image of the actual shape. Much of the learning tended to be wasteful since the verbal analysis was often arbitrary and unsystematic. Above all, the material available could not be conveniently used to give the students practice in repeated *acts of recognition.*

This was the training situation in 1942, when Dr. Samuel Renshaw of Ohio State University opened a school for spotters and navy instructors in aircraft recognition. Renshaw emphasized the necessity of learning plane shapes visually and conducted classes by repeatedly showing projected images on a screen, and requiring them to be identified. He added the feature of tachistoscopic exposure, or flash presentation, for which he had a great enthusiasm as a method of promoting visual learning. His procedures were widely publicized and were officially adopted by the Navy for training. They became known as the "Renshaw System," which was contrasted with the "WEFT System." The AAF adopted these methods, with some modifications, in early 1943 under the name of the "Flash System of Instant Recognition." Nearly all the Navy instructors and a great part of the AAF instructors were indoctrinated in it.

The basic assumptions and aims of this type of training were as follows:

a. That the principle aim of training was to develop instant recognition, i. e., an immediate identification of the stimulus object.

b. That this could be accomplished by increasingly more rapid exposure speeds during the course of training. It was assumed that the shorter were the exposure intervals at which pictures could be identified, the greater was the level of proficiency. Students were trained by the flash method until they could recognize

pictures at one seventy-fifth or even one one-hundredth of a second. This achievement was very impressive, and the demonstration that it could be attained by nearly anyone, with proper training, was primarily the fact which led to the adoption of the method officially.

c. That the method of flash-exposure forces the student to perceive the total form of the planes, and that they should be learned by "total form" rather than by "analysis of features" such as characterized the WEFT System.

d. That training in the perceiving of flash exposures, with material of the sort used by psychologists in span of attention experiments, improves the general efficiency of vision. More particularly, that the perceiving of increasing series of digits flashed on the screen widened the effective form field, i. e., increased the area of the retina over which form could be perceived.

According to the manual, *Instruction for the AAF Method of Recognition Training*, published by the AAF Training Aids Division on 15 September 1943, the level of proficiency to be striven for is the identification of planes at one seventy-fifth of a second. This proficiency can be attained by requiring instant recognition, by forcing the learning of total forms instead of features, and by improving the general effectiveness of visual perception. The improving of vision was said to be the basis of the entire course. Training in reading numbers with flash exposures, starting with four digits and gradually increasing to eight digits, was important for this purpose, since it would force the student to use portions of the eye that he would not use in everyday life and would increase the angle of the visual field over which he could distinguish form.

Equipment and Classroom Procedure. The materials employed in the flash system were primarily a film-slide projector and a slide kit containing as many different 2 x 2 inch photographic slides of the airplanes to be learned as were issued with the kit or as it was possible to procure. The projector, or "flash meter" as it came to be called, had a photographic shutter attachment which permitted variation in the exposure interval between one one-hundredth of a second on the one extreme and time exposures under the control of the instructor at the other. The slide kits were issued by the AAF Training Aids Divisions. They did not at the outset include many views of each of the planes to be taught, although it was recognized that the effectiveness of the training would be increased if a large number of slide views of each plane could be used.

The first course in aircraft recognition usually included about 40 different aircraft. This was the case in the Preflight School for aviation cadets at Santa Ana Army Air Base, where the experiments to be described were carried out. The course consisted of 30 one-hour sessions over a period of nine weeks. The classroom

procedure was to present two new planes each hour, showing pictures and standard silhouettes and discussing their characteristics, and then to conduct a slide review or corrected test, at flash speeds, of the planes already studied. This daily airplane review, as emphasized in the *Instructors Guide* for aircraft recognition, issued by the Training Command, was an important feature of the course. Ten to twenty planes were flashed serially and identified in writing by the class. After each flash exposure, the identifications were usually either confirmed or corrected immediately, the slide being shown again with a long exposure to permit the student to look at the picture while correcting his identification. This procedure had the advantages of providing frequent repetition, immediate correction or confirmation of the identifications made, and the maintenance of attention on the part of the students. Classes conducted in this way were usually interesting and the students were well motivated.

The use of digit slides for improving perception was also prescribed for each classroom hour, together with the use of "counter" slides, consisting of varying numbers of airplanes pictured at a distance, the requirement being to report the number of planes seen. The purpose of counter training, according to the proponents of the Renshaw System, was to accustom the student to the task of approximating the number of planes in a distant formation within the short time interval which combat conditions permitted him. It will be recognized as an adaptation of one type of the span of attention experiment. The utility of counter and digit training was regarded with suspicion by some recognition instructors, and they were unofficially omitted from classroom instruction in many schools well before they were officially eliminated from the prescribed course. Certain other procedures originally recommended, such as the use of a fixation point on the screen, the blurring of the image by off-focusing the projector or the cutting down of the brightness of the image to simulate "unfavorable conditions," and others, never were widely employed because of suspicion regarding their correspondence to real situations.

The practices described, and the measuring of proficiency in terms of the rapidity of the flashes which could be perceived, were the subject of much discussion and some controversy by specialists in aircraft recognition. They were believed by some to constitute a type of progressive teaching contrasting favorably with the rote-memorization methods too frequently characteristic of Army training schools, and by others to constitute something approximating the practices of a cult. The experiments to be described were performed against the background of this difference of opinion.

703313—47—9

The Reasons for Training in Aircraft Recognition

The aim of training in aircraft recognition within the AAF was to make possible the identification of friendly and hostile planes during combat. Identification was necessary primarily because decisions had to be made to fire or to withhold fire at other planes. Tragic mistakes in this decision were sometimes made in the excitement and confusion of air combat and of air-ground or air-sea operations. A secondary reason for identification was that in air combat the size of an airplane which is to be fired upon must be known in order to determine its range. The wingspan of the plane in relation to its projected angular size in the gunsight will give its range. In the case of the computing sight, this is done automatically. Wingspans have been memorized by aerial gunners, but in order to remember the wingspan and adjust for range the plane must first be identified.

Late in the war, the point was made that the gunners in a combat crew do not in actual fact fire or withhold fire on the basis of a visual identification of the shape of a plane, but on the basis of whether or not it attacks, i. e., its behavior. In the last analysis it is perfectly true that the final criterion of an enemy is whether he attacks you. This was particularly true for flexible gunners whose job is defense. Their practice was to shoot at any fighter plane which "pointed its nose" at the gunners plane, i. e., at any plane which entered upon an attack no matter what it looked like. It might, even if its shape were friendly, be a captured plane which was being used by the enemy. And in any case it is no better to be shot down by a friend in error than by an enemy. The gunners could hardly be blamed for such a practice. When a fighter plane does commit itself to an attack, it remains as a target and as a menace only a few seconds, during which time the gunner must continually adjust his aim and fire; it is only reasonable that during this brief interval he will have neither the time nor the inclination for recognition.

This practice of combat gunners did not, however, demonstrate the uselessness of learning to identify aircraft by shape, as was rather widely concluded. It demonstrated only the uselessness of recognizing aircraft with hairtrigger promptness *after they were committed to an attack*. The time for recognition is not during an attack, when it is already too late, but before the attack begins when the airplane in question is still out of range, i. e., at more than 1,000 yards. Recognition at that time permits the gunner to make preparations and plan his fire beforehand. If the plane is friendly in shape he can at least assume that attack is improbable and can direct his attention to scanning other regions of the sky. The kind of recognition that was required was early recognition rather than instant recognition.

The method of training which stressed instantaneous recognition was set up without knowledge of the actual military situations which fliers would have to face. No practical experience, it is true, was available since the situations had never before existed. It was assumed that only the briefest interval would intervene between seeing an enemy airplane and having to fire at it. But the assumption was not in accordance with the facts. The usual situation was the appearance of one or several aircraft in the sky at great distances. If the sky were being systematically scanned, they would be observed. They could be identified as soon as they got within recognition range. There was some discussion of what recognition range was, but it was unquestionably greater than the firing range. The emphasis in recognition training, consequently, should have been placed on accuracy of recognition at maximum distances rather than on speed of recognition. The kind of training which required the motionless fixating of the eyes on the center of a projection screen so as to see a flashed image was also not appropriate to the kind of observation required in searching the sky for enemy aircraft.

Aircraft Recognition as a Form of Perceptual Learning

There were in operation during World War II between 40 and 60 types of aircraft which might be encountered in combat operations of the AAF. The identification of all these planes in flight was the aim of recognition training. The number might have been cut in half by training for only those planes existing in a single theater of operations except for the fact that a flier had to be trained, at least for many months, for whatever theater he might be needed in. Military airplanes are relatively similar to one another in general shape and appearance. The same plane, moreover, has a very different shape when seen from the side, the front, and from below. The task of discriminating all of them at a distance and at any angle of view so that they can be identified and named is therefore a difficult visual performance. It has been suggested that learning planes is like learning faces.

A somewhat similar learning situation would be that encountered by a teacher faced with a new class of strange pupils. They look very much alike, yet there are differences. In the course of time these differences are discriminated; the features and outstanding peculiarities stand out as distinctive and the teacher learns to call each pupil by name, i. e., he learns to recognize them. The method by which he does so is repeated association of the visual appearances with names. When students who look alike are misnamed they are quick to correct him.

Another similar learning situation is that of learning to associate nonsense syllables with unfamiliar visual forms, an abstract

experiment performed in the psychological laboratory. It is known that the more similar the forms, the more difficult is the learning. It is also known that confusions occur, i. e., the misnaming of forms which look alike. These may be called generalized responses. As these errors are corrected and the correct responses are reinforced, the errors drop out and the responses become differentiated. The latency of the responses will decrease with practice. Conditioned response theory explains this process as one of differential reinforcement.[1] Presumably such a process of differential reinforcement of right and wrong naming responses would also be required to discriminate and identify such mutually similar shapes as military planes.

As differentiation of the naming responses progresses, it must be supposed that visual memory images develop of the forms (or of the students) and also become differentiated from one another. In the process, they become organized and acquire meaning. Studies of drawings made by subjects required to learn visual nonsense forms suggest that the insignificant features of these images tend to be eliminated, while the significant features tend to be exaggerated.[2] If it can be assumed that the development of differentiated memory images goes hand-in-hand with the development of differential naming responses, then the method of drawing aircraft should be a useful supplement, in training, to the method of naming them. This assumption can be verified experimentally.

As a basis for the experiments, it was therefore assumed that recognition training is essentially a kind of perceptual learning in which visual shapes not at the outset distinctive become capable of producing differential reactions. As repetition mounts the right reactions are progressively reinforced and the confusions drop out. The responses should become more immediate and more certain with practice, and it is this factor which should be expected to produce prompt or "instant" recognition. With regard to what the student experiences, his perceptions may be assumed to acquire meaning and the shapes will become differentiated, fully organized or precise. They are also, surely, unitary perceptions, i. e., planes, and in that sense are seen and learned as "total forms."

The Problem of Materials for Classroom Instruction

The problem of what kind of materials should be used to present to the students the airplanes they were required to recognize was an important aspect of the effort to improve recognition training. Advantages were claimed for a number of methods of reproducing the aircraft (photographs, slides, models, motion pictures) and

[1] Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, 47, 196-229.

[2] Woodworth, R. S. *Experimental Psychology*, New York, Holt, 1938, Ch. 4.

many methods of setting up the task of recognition (moving models, miniature theaters, booklets, packs of cards, posters, displays, projected shadows, projected pictures). A variety of "trainers" purporting to imitate the requirements of the real situation had to be evaluated by the AAF Training Aids Division in this field as in nearly all other fields of training. The problem was one aspect of the general problem of "synthetic training", i. e., training which is artificially reproduced instead of naturally encountered. Although no formal experiments were undertaken by the Film Unit on this problem, it was involved, directly or indirectly, in a number of the research reports submitted.

Under actual combat conditions, planes had to be recognized at various distances, under a variety of haze and lighting conditions, and in any of a theoretically infinite number of attitudes. Planes would seldom appear in one of the standard silhouette attitudes, and frequently were so far away that only the most outstanding details of their shapes were discriminable. As a first criterion, therefore, if training is intended to produce proficiency in recognition under such conditions, the type of material chosen for training should be capable of representing planes in a variety of attitudes and distances. It was maintained by some, also, that representation should be of three-dimensional forms, using stereoscopic slides and polarizing eyeglasses. However, at the usual distances at which planes are recognized, the cues to the perception of solidity are absent. The advantage of three-dimensional forms such as models seems to reside in the possibilities they offer for the presentation of a large number of different attitudes for each plane, rather than in the fact that they are solid.

A second criterion, proposed for presentation materials, was convenience for the instructor. They should be of such a nature as to make constant review an integral part of the course. To some extent this was a matter of number, rather than type of materials. For example, if photographs were used, it was essential that the instructor be provided with enough photographs of the same plane to enable him to review frequently without teaching the recognition of a particular photograph.

The four main types of materials met these criteria as follows:

a. *Photographs.* Real plane photographs made possible the presentation of a variety of attitudes, distances, and lighting conditions. Their usefulness depended upon the number of photographs of each plane. Projection of photographs on a screen was the most convenient method of presentation to a group, but involved the use of a projector which is relatively cumbersome.

b. *Slide Photographs.* These were the most widely-used materials in training for aircraft recognition. It was ultimately found possible to provide a large number of slide views of each plane

taught. The possession of these slides and a convenient apparatus for exposing them on a screen, the slide-projector or "flashmeter," undoubtedly made possible the frequent reviews which were such a large portion of the aircraft recognition course.

c. *Models.* Models could be used most efficiently only when some method was provided for repeated presentations in various attitudes. One such method was the "shadowgraph" technique, in which the shadow of a model, manipulated into various positions by the instructor, was thrown on a translucent screen and viewed by the students. Although reviews were more troublesome to conduct under these conditions, in the hands of a skilled instructor the method provided a means of showing planes in a great variety of changing attitudes, with contours as they would naturally appear at great distances. Other methods of using models were useful, particularly one which displayed a group of planes together on numbered hooks, thus permitting a quiz or examination with reference to the numbers. The mere passive *displaying* of models either in or out of the classroom should not produce effective learning unless it was required that they be actively identified. The same rule applied to the displaying of posters.

d. *Motion Pictures.* Motion picture shots of real planes in flight made possible representation of planes in a great variety of attitudes, distances, and lighting and background conditions. Above all, they would yield an impression of a single object with different appearances in different attitudes, and would thus assist in the process of combining different views into a single perception—a process which was difficult when isolated slide views were the only mode of presentation. They were relatively easy to project in class. Films of the "testcraft" type guaranteed a convenient and effective review; the more so if the principle of reinforcement of learning had been considered in constructing the films. Since the list of planes to be taught was altered frequently, motion pictures suffered somewhat from the fact that they became out of date. However, films of this type could in theory be kept up to date by splicing. The chief drawback of training films as a medium of instruction is the loss of relationship between the student and his instructor. Since a response is not demanded of the student by the usual type of training film produced for use in the aircraft recognition classroom, and since his questions necessarily go unanswered, he may actually make no response, in which event learning will suffer.

Any of these methods could be effective, provided that certain conditions were met. It was chiefly necessary to provide a sufficient number and variety of views and to make possible a convenient method of repeated review. The demonstrated convenience

of slide pictures, their flexibility, small bulk, and ease of handling, was the consideration most in their favor.

Drawing the Aircraft. The reactions which the students were required to make to any of these types of presentation (when they were required) were usually acts of identification, i. e., naming. The name or designation of the plane and its memorized wingspan had to be written or spoken. Other information was sometimes required but not often. Another type of response, however, was enthusiastically advocated by some instructors, namely that of drawing the aircraft from memory, in a specified view, or copy- its outline freehand from a picture or silhouette. This response is clearly not like that required in combat recognition. Nevertheless, the ability to draw a shape should be indication that it has been learned and differentiated from other shapes. Whether proficiency in drawing planes is correlated with proficiency in identifying them was a question which required an empirical answer.

The Questions for Experimental Investigation

The following questions appeared to be the ones on which experimental evidence was required if the most effective methods for recognition training were to be objectively determined. The first three listed, concerning methods in use under the orthodox flash system, were asked in the original letter from the Assistant Chief of Air Staff for Training to the Air Surgeon requesting experimental evidence.

1. Are rapid-flash speeds of practical value in the learning of aircraft recognition?

2. Is instruction emphasizing the total form of the aircraft more effective than instruction which emphasizes an analysis of the features of the shape to be identified?

3. Does training in reading digits at flash exposures, or in estimating the number of spots in a flashed presentation (counter training) improve proficiency in aircraft recognition? Does such training improve the general efficiency of vision?

4. What are the characteristics of the visual memory images of aircraft as shown by drawings, and is the ability to draw the shapes of aircraft indicative of the ability to identify them by name?

5. How significant is the method of differential reinforcement of responses, as practiced in plane reviews, for the learning of aircraft,

6. In view of the tendency to confuse similar planes, should similar planes be introduced together for contrast or separately to minimize initial interference?

7. What features of each plane are significant for recognition, i. e., what are the identifying or characteristic features which

distinguish it from others? Can a conceptual organization of the various shapes be constructed so as to promote differential learning?

8. Can the range of aircraft be represented in pictures and what is the relation of apparent range to the distance at which the picture is viewed?

9. What is the recognition range of different airplanes against the sky in relation to the firing range?

An experiment was devoted to each of these nine questions. They will be reported in the order listed. All results were obtained with aviation cadets in the Preflight Schools at Santa Ana Army Air Base between December 1943 and January 1945.

THE EFFECTIVENESS OF CERTAIN ASPECTS OF THE FLASH SYSTEM OF INSTRUCTION

The Efficiency of Rapid Flash Speeds

Introduction. Once recognition training had been set up in the AAF with slide projectors, slide kits, and the practice of daily reviews, the practical question remained whether the feature of flash training actually facilitated the learning. The projectors could be used for tachistoscopic exposures, but they could also be conveniently used for exposures of any desired duration.

Among a number of advantages claimed for the flash method by instructors trained by Dr. Renshaw only two were actually relevant to whether flash exposures were necessary for efficient learning of the planes: that flash presentation produces superior learning because it forces the student to learn "total forms," and that flash presentation facilitates learning because it mobilizes attention in the classroom and provides a high degree of motivation and interest. The second of these arguments was fairly convincing, especially to anyone who had seen the method in use; the first argument was a theoretical one and was at least debatable. An experiment comparing the effectiveness of split-second exposures and slow exposures in training could, however, settle the question without resort to arguments.

It should be pointed out that the arguments stated above were formulated to contrast the flash method with a procedure which consisted chiefly of looking at and memorizing pictures—the so-called "WEFT System." The experiment on rapid *vs.* slow exposures, retaining in both cases the feature of systematic drill in the identification of slides, had not been tried. The method of tachistoscopic viewing, at exposures too rapid to permit the moving of the eyes over the picture, was only one feature of the flash system in practice; the other was systematic drill. The potential benefit for learning of the latter feature was lost sight of, how-

ever, by enthusiasts for the former. In the experiment to be described, flash presentation was compared with presentation at one-second exposures. The frequency of exposures and the amount of drill was the same in both cases.

Two other advantages were claimed for the flash method of training, having to do primarily with its appropriateness for combat conditions. The first of these was the assumption that flash exposures represent the kind of perception in glimpses required of fliers in aerial warfare. The actual facts, as we have seen in an earlier section of this chapter, showed the assumption to be in error. The second was the assumption that split-second presentations promote prompt identification, i. e., instant recognition. The earlier discussion concluded that to put a premium on an immediate or instant reaction in combat identification was a mistake. If this conclusion is correct, the advantage claimed for instant recognition fails to be valid. But the proposition may be challenged as such. On purely psychological grounds, the inference that instantaneous exposures are necessarily connected with or produce instant responses is unjustified by any known experimental facts. It has, to be sure, a certain common sense plausibility, but it is contradicted by the evidence from experiments on reaction-time which show that the latency of a response is not appreciably affected by the duration of a visual stimulus. In the case of very brief, and therefore "improverished," stimuli the tendency is for the reaction to be, if anything, slower.[3]

The Method of the Experiment. Three groups of four classes each were used in the experiment, each group consisting of about 160 students of aircraft recognition. Four instructors carried out the teaching in the experimental classes, each instructor having one class in each of the three experimental groups. Hence differences in the quality of instruction were distributed in an equal fashion throughout the three groups.

In one experimental group, training on all slides was carried out at an exposure speed of one second throughout the course. In the second experimental group, training was carried out at a moderate flash speed of one-tenth second throughout, except for the first few hours of instruction in which one-fifth second exposures were employed. The third experimental group was trained at an exposure speed of one-fiftieth second, reaching this speed during a two-week period in which the exposures were gradually increased from one-fifth to one-fiftieth of a second. These three groups will be referred to as Slow-Trained, Intermediate-Trained, and Fast-Trained, for the remainder of this report. The Slow-Trained Group never saw slides presented at split-second intervals; a one-second exposure is not a "flash speed." The same num-

[3] Woodworth, R. S., *Experimental Psychology*. New York: Holt, 1938, ch. 14.

ber of slides and the same type of slides were used in each of the experimental groups. The instructors devoted the same amount of time to "airplane review" in all cases.

The Criterion of Proficiency. The proficiency of the three groups was measured by the preliminary form of the motion picture test described in chapter 6, the Aircraft Recognition Proficiency Test. This criterion was not dependent on skill in perceiving flash exposures, which was a question at issue. This criterion was assumed to be the most realistic one. In order to compare the learning of the three groups by all possible criteria, however, a three-part final slide examination was also made up especially for the use of the experimental classes, consisting of 60 slides. The examination consisted of three sets of 20 slides labeled A, B and C, each set being of approximately equal average difficulty. One-third of each group of students viewed the slides of set A at one-second exposures, the slides of set B at one-tenth second exposures, and the slides of set C at one-fiftieth second exposures. Another third viewed set A at one-fiftieth second, set B at one second and set C at one-tenth second. The final third viewed set A at one-tenth second, set B at one-fiftieth second and set C at one second. Differences in difficulty between the three sets of slides making up the examination were therefore counterbalanced. This procedure insured that any differences which might appear between the three parts of the examination could be attributed solely to differences in exposure speed rather than to possible differences in the difficulty of slides themselves.

Results. The results obtained make possible a comparision of the proficiency exhibited between the Fast-Trained Group, the Intermediate Group, and the Slow-Trained Group at the end of the standard 30-hour course in aircraft recognition. The mean score made by each group on the motion picture test together with its standard deviation was as follows:

Slow-trained		*Intermediate-trained*		*Fast-trained*	
M	*SD*	*M*	*SD*	*M*	*SD*
63.4	13.6	61.9	12.2	62.2	13.8

The differences between mean scores are small and the critical ratios of these differences are not significant. They were: between Slow-Trained and Intermediate-Trained, 1.06; between Slow-Trained and Fast-Trained, 0.85; and between Intermediate-Trained and Fast-Trained, 0.17. There is evidently no advantage so far as the motion picture test shows, in training with rapid-flash speeds. Although the motion picture test showed no advantage from the use of flash training, it might be expected that an examination carried out at flash exposures would show such an advantage. The results from the three-part slide examination are given in table 7.1. The mean scores represent number of slides

TABLE 7.1.—*Proficiency in aircraft recognition of three experimental groups as measured by slide examinations at different exposure speeds*

Group	N	Test at 1 sec.		Test at 1/10 sec.		Test at 1/50 sec.	
		M	SD	M	SD	M	SD
Slow-trained	178	16.7	2.2	14.3	3.0	13.0	2.9
Intermediate-trained	167	16.7	2.3	14.3	2.8	13.6	2.9
Fast-trained	177	17.0	2.4	14.0	2.8	14.1	3.3

correctly identified out of 20. Running down the columns of the table, it may be seen that the one-second test and the one-tenth second test do not show any evidence of superior proficiency for any experimental group. None of the differences between mean scores in the first two columns of the table are significant. The critical ratios vary between 0.2 and 1.4. The only significant difference is between the Slow-Trained Group and the Fast-Trained Group when both are tested at one-fiftieth second, the critical ratio here being 3.27. It is probable that this difference may be attributed to the fact that the Slow-Trained Group had no opportunity of learning the specific viewing habits required for seeing slides at one-fiftieth second. By all other criteria the proficiency of the Slow-Trained Group was not inferior to that of the Fast-Trained Group. In general, proficiency was equivalent whether training was conducted at one second, one-tenth second, or one-fiftieth.

The data of table 7.1 indicate that the highest scores were obtained by all three groups on the test at one second, regardless of the speed at which training was conducted. The differences reading across the table are all statistically significant with the exception of those between the Intermediate and Fast-Trained Groups at one-tenth second as compared with one-fiftieth second. It is evidently easier to recognize slides at an exposure speed of one second than it is at split-second intervals. This result is of some interest in view of the claim sometimes made by the proponents of flash training that for properly-trained students it is actually easier to identify slides at split-second intervals than at longer intervals.

Conclusions. No differences in recognition proficiency were discovered when training with flash slides was conducted at exposures of one second, one-tenth second, and one-fiftieth second.

The evidence obtained indicates that slides shown for one second are easier to see than those shown for either one-tenth or one-fiftieth second, in that fewer errors are made in a slide test of recognition given at one-second exposure speed, regardless of the type of training of the individuals tested.

It is concluded that there is no practical value in training individuals to recognize slides of aircraft at rapid flash speeds. The use of exposure speeds as long as one second is probably to be preferred when the slide content is difficult. According to the evidence obtained, efficiency of training with one-second expo-

sures is equal to that which is attained with split-second intervals. Longer exposures than one second were not further investigated. It was reasoned that exposure intervals considerably greater than a second would have the effect of decreasing the number of slides which could be reviewed in a single class period. The possibility of a slight but significant superiority of two- or three-second exposures exists.

The Relative Importance of Emphasizing the Total Forms or the Features of the Airplanes to be Identified

Introduction. One of the reasons given for the employment of rapid-flash speeds in the presentation of aircraft slides was that these rapid exposures were said to discourage the analysis of the features of airplanes by the student, and thus to bring about more rapid identification. It was assumed that in making such an analysis of features as was called for in the so-called WEFT system, the student failed to learn the appearance of the overall form of the plane, which he must know in order to recognize it rapidly. The flash method of training aimed to produce perceptions of the "total forms" of aircraft. The "learning of total forms" became something approaching a slogan among instructors in the recognition course and its implications were widely discussed. By some it was interpreted to mean that analysis of the features or characteristics of aircraft shapes in the classroom was "wefting" and was something to be avoided. It was said to lead to "disjunctive seeing" and inefficient perception. When, however, it became a matter of actually discussing the airplanes with students, the instructors found it difficult to do so except in terms of these features. Other instructors, probably in the majority, argued that the distinguishing characteristics of the shapes to be learned were an important aid to learning and therefore analyzed them freely, using for the purpose a more-or-less unsystematic repertory of terms, such as "squared wing tips," "barrel-shaped fuselage," "bubble canopy," and "high, faired vertical."

The Film Unit did not participate in this controversy, believing that the issue being debated was abstract and often confused. The most reasonable psychological position seemed to be that there could be good or bad analysis of the characteristics of a plane, and that visual-form perception was "total" by its very nature. The hypothesis that there could be two types of form perception, "disjunctive" on the one hand and "total" on the other, was regarded as unproved.

In view of the possibility of such types of perception and of the actual difference of opinion on how to conduct classes, an experiment was designed to test the effectiveness of two methods of classroom presentation of aircraft. One involved emphasis on

learning the total forms without reference to their features. In the other a pre-standardized set of distinctive features was emphasized.

Method. Six classes in aircraft recognition in the Preflight School (Pilot), Santa Ana Army Air Base, were employed in the experiment. These classes, containing about 30 men each, were divided into two groups which were equalized for recognition proficiency on the basis of scores on a pretest of 20 slides of relatively familiar American aircraft. Three instructors taught these classes, each instructor having one class in each group, so that differences in teaching ability could not affect the results.

Both groups of classes were taught the same planes in the same order, by the usual flash-slide method. Equal numbers of slides and equal amounts of time were used in the reviews of each group. Slides were reviewed at flash speeds of one-tenth second. However, during the presentation of each plane, the method of instruction differed in the two groups. In one group, instruction was given on only the total form of each plane, and no mention whatever was made of features such as shapes of wings, nose, or tail. The students were encouraged to become acquainted with the plane as a total configuration, and to learn to recognize it by its overall form. In the second group, a standard set of distinctive features, agreed upon by the instructors, was emphasized in the case of each plane presented. The students in this group were encouraged to learn these features for each plane, and to identity the plane by means of them. At the end of the thirty-hour training period, both groups of classes were tested by means of a slide examination composed of 45 slides showing the 40 planes presented during the course. In addition, they were tested with a copy of the Aircraft Recognition Proficiency Test, Preflight Level, a motion picture test covering the planes which had been taught. The pretest on which the groups were matched correlated .47 with the Slide Examination, and .68 with the Motion Picture Test.

It should be understood that the control of formal classroom instruction in the manner described does not make possible complete control of experimental conditions. In the present case, for example, the students could not be prevented from reading and studying training manuals and spotters' guides outside of class. Therefore, the students of the "total form" classes had the opportunity of studying the features of planes if they so desired, even though these were not emphasized in class. The actual difference in method of learning in the two groups of classes may have been a relatively small one. The results provide an answer only to the practical problem of whether to recommend the changing of certain aspects of formal instruction in the classroom. It would be a mistake to suppose that they provide a conclusive answer

to the theoretical problem of how the learning of perceptual forms takes place.

Results. The average number of planes correctly identified out of 45 on the final slide examination was 40.0 (SD of 3.96; N of 87) for the group which had been instructed by emphasis on total form of airplanes. For the group which was instructed by giving emphasis to features, the average score on the same examination was 41.2 (SD of 3.57; N of 98). The critical ratio of the difference between these two averages, corrected to allow for the reduction in differences arising by sampling, is 2.44. This figure indicates that a difference of this size could be expected to occur by chance between 1 and 2 times out of 100. A probability of chance occurrence as high as this indicates a difference which is considered on the borderline of statistical significance. Since the group instructed by the "total form" method did have some opportunity to learn plane features outside the classroom, this extra study may have had the effect of reducing the obtained difference in score. The results suggest that there is a slight disadvantage in effectiveness for recognition training for the exclusive emphasis on "total form" when proficiency is measured by means of the final slide examination.

The group of men who had been taught recognition by total form made an average score of 61.03 (SD of 11.4; N of 87) on the Aircraft Recognition Proficiency Test. The group whose instruction had emphasized features obtained an average score of 62.9 (SD of 12.0; N of 93). The difference between these average scores has a critical ratio of 1.49, indicating that such a difference might occur by change 14 times out of every 100 similar experiments. Here again the advantage of feature-emphasis appears to be slight, but may be interpreted in the manner indicated in the previous paragraph.

Two principal observations made by the instructors who took part in the experiment are of some bearing on the question of the effectiveness of the two methods under consideration. The impression was obtained by all three of the instructors, at about the time the course was two-thirds completed, that the group taught by emphasis on total form was definitely "slipping" in comparison with the other group. The second observation was that a single question was insistently and repeatedly asked by the cadets in the group taught by the method which emphasized total form. This question was "How can I distinguish between this plane and the one which resembles it closely (e. g., the C–46 and C–47)?" The students in this group apparently felt no need for emphasis on features, except *insofar as these features serve to distinguish one plane from another.* Showing the two confusable planes on the screen at the same time was not a satisfactory

answer; the students wanted to know in descriptive terms what the differentiating features of the two planes were.

Conclusions. The results have shown that the method of instruction which emphasizes the "total form" of airplanes without mention of features is, if anything, slightly less efficient in training for aircraft recognition than a method which emphasizes features. The group trained by the former method made lower scores on both the final slide examinations and the Aircraft Recognition Proficiency Test, Preflight Level, than did the group trained by the method which emphasized airplane features. Although differences between the average scores made by these groups on both these tests are not significant, they indicate a probability of the relative ineffectiveness of an exclusive emphasis on "total form." This result is in contradiction to some claims made in connection with the "flash system." The experiment fails to verify the existence of two kinds of visual perception, a superior total type and an inferior disjunctive type.

The reactions of the students to the method of instruction which emphasized "total form" indicated that they felt a need to know those features which distinguish one plane from other similar, and therefore confusable, planes. These reactions suggest that it is desirable for effective training in aircraft recognition to emphasize the features of planes, particularly those which distinguish similar planes from each other. Features which were being emphasized were not necessarily chosen with this idea in mind. In all probability, much time which was devoted to emphasis on *unimportant* features during the presentation of planes, could more profitably have been devoted to a discussion of *distinguishing* features of confusable planes.

The Value of Supplementary Training in Reading Digits and Counting Spots with Flash Presentation

Introduction. One of the techniques employed in the flash system was the use of slides showing series of from 4 to 10 digits (digit slides) and slides showing groups of from 3 to 30 planes (counter slides), which were presented at increasingly rapid exposure speeds. The students were required to reproduce the series of digits and to estimate the number of planes shown in the counter slides, the interval being too short for actual counting. Digit training was claimed by its proponents to increase the general efficiency of perception, and to "widen the angle of vision." Counter training was valuable in developing the "perception of numerousness." Both made it possible for the student eventually to see and recognize planes exposed for as short as one seventy-fifth second. Experience with these materials by instructors of aircraft recognition left little doubt that improvement does occur

in reading digits and in estimating numbers. The question was whether such training transferred to the recognition of aircraft.

Except for the fact that both involved the use of rapid flash speeds, there was, after all, little resemblance between the activities involved in counter and digit training and those involved in recognizing airplanes. If counter and digit training developed *specific* habits related to these activities alone, there was no reason to suppose that recognition proficiency would be influenced by such training. If, on the other hand, counter and digit training led to an increase in general perceptual efficiency, this training might be expected to influence recognition proficiency insofar as the latter is affected by perceptual factors.

An experiment was therefore undertaken to determine, in the actual classroom situation, the effect of training with counter and digit slides on, first, proficiency in aircraft recognition, and, second, on the efficiency of a different kind of perception requiring what might be considered a wide span of attention.

Method. Six classes of cadets in the AAF Preflight School (Pilot) at Santa Ana Army Air Base, were used in the experiment. These were divided into two groups of approximately 100 cadets each. The two groups were equated for initial knowledge of airplanes on the basis of scores on a pretest containing 20 slides of United States airplanes shown at an exposure speed of one second.

During the first twelve minutes of each of 10 class hours, one group of cadets was given digit and counter training. Twelve digit-slides and six counter-slides were given each day. The range of digit series shown was gradually increased during the training period from 4 to 9, and the range of plane groups on the counter slides from 2 to 14. The speed of exposure of these slides was also gradually increased from one-fifth second to one twenty-fifth second. Each instructor used the same digit and counter slides, and introduced the gradual increases of numbers and flash speeds at the same points in training. The cadets in these classes were instructed as to the importance of this training and its influence on proficiency in aircraft recognition, in order that motivation might be maintained at a high level. The instructors made a genuine attempt to have their classes reach as high a level of proficiency as possible in this kind of performance by the end of the 10-hour training period. The second group of trainees received no counter and digit training during this 10-hour period. The first twelve minutes of each class hour were occupied by lectures and discussions of such topics as Organization of the AAF, Branches and Functions of Military Aviation, etc. Thus this second group received the same amount of instruction in actual plane recognition as did the counter- and digit-trained group.

Both groups were instructed on the same planes, and spent the same length of time on review of these planes.

The fact should be emphasized that the time devoted to counter and digit training in the experimental classes was one-quarter of the total available class time for the duration of the experiment. This amount was three times greater than that recommended in official instructors' guides, and was about as much as could be given practically without seriously interferring with plane-recognition instruction. It was not as much training as was recommended by Dr. Renshaw.

Proficiency in Reading Digits and Counters. At the conclusion of the 10-hour training period, the two groups were tested by means of an examination containing 12 digit slides with number series ranging from 5 to 9, and 8 counter slides with the number of planes shown ranging from 3 to 12. The slides were exposed at a speed of one-tenth second. The score on the digit test was the total number of digits contained in all the series which were reproduced with complete accuracy.. The score on the counter test was the total number of objects on all the slides for which the numbers of objects were correctly reported. Since the counter and digit training was supposed to develop the ability to perceive slides correctly at rapid exposure speeds, as well as a greater *number* of both digits and objects, it was believed that these scores would represent the required proficiency better than the more conventional measures of digit span or "span of apprehension." The correlation between the digit test and the pretest for initial knowledge of airplanes (by means of which the groups were matched) was .03; that between the counter test and the pretest was .23.

Proficiency in Aircraft Recognition. Both groups were tested at the end of the 10-hour training period by means of a slide examination containing 30 slides of the United States planes which had been covered in the course. The slides of this examination were shown at an exposure speed of one-tenth second. The correlation between this test and the recognition pretest was .57.

The proficiency of these two groups was also tested at the conclusion of the training period by means of a motion picture test composed of views of airplanes in flight, the Aircraft Recognition Proficiency Test, Preliminary Form. The cadets were tested only on the views of the United States planes taught in the course. There were 52 of these views. The correlation between this test and the pretest of recognition was .59.

Increase in General Ability to Perceive or Attend. The two groups were tested at the beginning and again at the end of the training period with the Flexibility of Attention Test (CP411E), a motion picture test which measures a certain type of perceptual

efficiency. It was described in detail in chapter 5. This film shows a set of 5 dials, lettered A, B, C, D, and E, the indicators of which vary continuously in a random fashion throughout each trial. The individual is required to record the letters of all the needles which reach certain blackened portions on the dials during each trial. The test demands constant attention on the part of the individual being tested, and samples the ability to perceive a number of events taking place in different portions of the screen at the same time, i. e., dispersed attention. It may therefore be considered as one of several possible measures of perceptual efficiency. Moreover, this test presents to the student a perceptual task which could be supposed to have elements similar to those of counter and digit training; i. e., the material is projected on a screen, it is spread out across the screen, and the whole of the projected material must be attended to. It appeared to be the sort of perceptual task to which *transfer* of counter and digit training would most easily take place, if such transfer occurred at all.

Some increase in score on the Flexibility of Attention Test when administered a second time was, of course, to be expected. The effect of counter and digit training on the performance tested by this film was measured by comparing the amount of improvement in scores of the trained and untrained groups. The initial scores of the two groups on this test were closely similar. The correlation between this "improvement score" and the recognition pretest was -.02.

Results of Training on Counter and Digit Proficiency. The average scores made by the trained and untrained groups on the digit test of 12 slides and on the counter test of 8 slides are given in table 7.2.

TABLE 7.2.—*Average scores on digit and counter tests*

	Trained group (N 102)		Untrained group (N 103)	
	M	*SD*	*M*	*SD*
Digit test	31.40	12.62	24.20	10.72
Counter test	30.44	8.32	28.36	9.02

These results show that the trained group made higher average scores on both these tests than did the untrained group. On the digit test, the critical ratio of the difference between these averages, corrected for the use of matched groups, is 6.17. The difference is a highly significant one. In the case of the counter test, the critical ratio of the difference between the averages of matched groups is 2.50, which indicates that a difference of this size could occur by chance 1.2 times in 100. The training may be said to have clearly improved the performance of the trained group over that of the untrained group with respect to digit lists and, to a lesser degree, with respect to counter material. For the purposes of the experiment it is sufficient to note that these results estab-

lished the fact that practice on the digit and counter slides produced an improvement in performance.

Results of Training on Recognition Proficiency. The average scores made by the trained group and by the untrained group on the slide examination and on the motion picture test are given in table 7.3. These results show that both groups made approximately the same scores on both these measures of recognition. The differences between the means are not significant ones, the critical ratios (corrected to allow for the use of matched groups) being .31 and .10. In interpreting these results, it should be kept in mind that the motion picture test, particularly, provides a sensitive measure of proficiency. The range of scores obtained was from 17 to 52, so that the test might be expected to discriminate successfully any differences in proficiency between the two groups. The proficiency scores of the trained and untrained groups indicate clearly that the digit and counter training did not improve proficiency in aircraft recognition. Since the special training given during the experiment occupied about one fourth of the effective teaching time, it may be concluded that more aircraft recognition could have been learned if the time had been utilized in the identifying of aircraft instead of in the perceiving of digits and counters.

TABLE 7.3.—*Average scores for recognition proficiency*

	Trained group (N = 102)		Untrained group (N = 102)	
	M	SD	M	SD
Slide examination	26.29	3.02	27.26	2.51
Motion picture test	38.75	6.16	38.77	5.20

Results of Training on Perceptual Efficiency. The amount of improvement in score on the Flexibility of Attention Test made by both groups is given in table 7.4.

TABLE 7.4.—*Improvement on flexibility of attention test*

Trained group (N 94)		Untrained group (N 92)	
M	SD	M	SD
7.21	8.45	6.95	8.15

As this table indicates, the average amount of improvement was not different for the two groups. The critical ratio of the difference between these two improvement scores of matched groups is .21, indicating that the difference is not significant. These results demonstrate that the training produced by practice on counter and digit slides did not generalize to the extent of improving performance on this particular test of perceptual efficiency. This fact casts considerable doubt upon the hypothesis that digit and counter training is capable of improving "general perceptual efficiency."

Conclusions. The ability to reproduce digit-series and to apprehend the number of planes in a group (counters)' was found to be

efficiency. It was described in detail in chapter 5. This film shows a set of 5 dials, lettered A, B, C, D, and E, the indicators of which vary continuously in a random fashion throughout each trial. The individual is required to record the letters of all the needles which reach certain blackened portions on the dials during each trial. The test demands constant attention on the part of the individual being tested, and samples the ability to perceive a number of events taking place in different portions of the screen at the same time, i. e., dispersed attention. It may therefore be considered as one of several possible measures of perceptual efficiency. Moreover, this test presents to the student a perceptual task which could be supposed to have elements similar to those of counter and digit training; i. e., the material is projected on a screen, it is spread out across the screen, and the whole of the projected material must be attended to. It appeared to be the sort of perceptual task to which *transfer* of counter and digit training would most easily take place, if such transfer occurred at all.

Some increase in score on the Flexibility of Attention Test when administered a second time was, of course, to be expected. The effect of counter and digit training on the performance tested by this film was measured by comparing the amount of improvement in scores of the trained and untrained groups. The initial scores of the two groups on this test were closely similar. The correlation between this "improvement score" and the recognition pretest was –.02.

Results of Training on Counter and Digit Proficiency. The average scores made by the trained and untrained groups on the digit test of 12 slides and on the counter test of 8 slides are given in table 7.2.

TABLE 7.2.—*Average scores on digit and counter tests*

	Trained group (N 102)		*Untrained group (N 103)*	
	M	*SD*	*M*	*SD*
Digit test	31.40	12.62	24.20	10.72
Counter test	30.44	8.32	28.36	9.02

These results show that the trained group made higher average scores on both these tests than did the untrained group. On the digit test, the critical ratio of the difference between these averages, corrected for the use of matched groups, is 6.17. The difference is a highly significant one. In the case of the counter test, the critical ratio of the difference between the averages of matched groups is 2.50, which indicates that a difference of this size could occur by chance 1.2 times in 100. The training may be said to have clearly improved the performance of the trained group over that of the untrained group with respect to digit lists and, to a lesser degree, with respect to counter material. For the purposes of the experiment it is sufficient to note that these results estab-

lished the fact that practice on the digit and counter slides produced an improvement in performance.

Results of Training on Recognition Proficiency. The average scores made by the trained group and by the untrained group on the slide examination and on the motion picture test are given in table 7.3. These results show that both groups made approximately the same scores on both these measures of recognition. The differences between the means are not significant ones, the critical ratios (corrected to allow for the use of matched groups) being .31 and .10. In interpreting these results, it should be kept in mind that the motion picture test, particularly, provides a sensitive measure of proficiency. The range of scores obtained was from 17 to 52, so that the test might be expected to discriminate successfully any differences in proficiency between the two groups. The proficiency scores of the trained and untrained groups indicate clearly that the digit and counter training did not improve proficiency in aircraft recognition. Since the special training given during the experiment occupied about one fourth of the effective teaching time, it may be concluded that more aircraft recognition could have been learned if the time had been utilized in the identifying of aircraft instead of in the perceiving of digits and counters.

TABLE 7.3.—*Average scores for recognition proficiency*

	Trained group (N = 102)		*Untrained group (N = 102)*	
	M	*SD*	*M*	*SD*
Slide examination	26.29	3.02	27.26	2.51
Motion picture test	38.75	6.16	38.77	5.20

Results of Training on Perceptual Efficiency. The amount of improvement in score on the Flexibility of Attention Test made by both groups is given in table 7.4.

TABLE 7.4.—*Improvement on flexibility of attention test*

Trained group (N 94)		*Untrained group (N 98)*	
M	*SD*	*M*	*SD*
7.21	8.45	6.95	8.15

As this table indicates, the average amount of improvement was not different for the two groups. The critical ratio of the difference between these two improvement scores of matched groups is .21, indicating that the difference is not significant. These results demonstrate that the training produced by practice on counter and digit slides did not generalize to the extent of improving performance on this particular test of perceptual efficiency. This fact casts considerable doubt upon the hypothesis that digit and counter training is capable of improving "general perceptual efficiency."

Conclusions. The ability to reproduce digit-series and to apprehend the number of planes in a group (counters)' was found to be

improved in the group which received training in these activities. Proficiency in aircraft recognition, as measured both by a slide examination and by the Aircraft Recognition Proficiency Test, *was found to be no greater in the group given counter and digit training than in an equivalent untrained group.* No significant difference in amount of improvement in score on a perceptual test, Flexibility of Attention, was found between the group given counter and digit training and the untrained group. This result calls into question the hypothesis that such training "improves the general efficiency of perception."

THE MOST EFFECTIVE METHODS OF LEARNING

The first questions put to experiment by the Psychological Test Film Unit concerned the debatable value of methods derived from the Renshaw system. They have been treated in the preceding section. Other experiments could be formulated which might be expected to lead to positive rather than merely negative recommendations, and these are reported next.

A Study of the Remembered Shapes of Aircraft as Revealed by Drawings and by Composites Constructed From Them

Introduction. A theory of learning to identify aircraft could be formulated primarily in terms of differential responses or primarily in terms of differential images or perceptions. An experimenter in this field might well be inclined to reject one formulation in favor of the other, his choice depending on his theoretical bias. The relation of learning to perceptual experience on the one hand and overt behavior on the other is a matter of controversy in present day psychology. There is a possibility, however, that the acquiring of differential responses and the acquiring of differential images are only two aspects of the same process. If so, they should be correlated. The aircraft recognition course provided an opportunity to measure this correlation.

The theory of the development of differential responses to stimuli has been worked out in some detail. A corresponding theory of the development of differential memory images is much less clear. Gestalt psychologists have attempted such a theory[4], but the experimental evidence on which to base it is relatively difficult to obtain. The best evidence comes from experiments on the learning of visual forms as evidenced by drawings, which are not easy to deal with experimentally and quantitatively. Nevertheless, both a theory of this type and the method of drawing visual forms are applicable to the aircraft recognition problem. The assumption that students trained to make differential naming

[4] Koffka, K., *Gestalt Psychology*, New York: Harcourt-Brace, 1935, Ch. 10 and 11.

responses acquire differential memory images can be verified by requiring them to make drawings of the aircraft at the end of their training. An image may be defined as a visual process which mediates drawing from memory. Predictions regarding the nature of these differential images, as revealed by the drawings, can then be checked.

The practical value of studying the remembered shapes of aircraft as represented in drawings by students would be, above all, to see what these shapes are like. The learning of "total forms" or the learning of "features" are vague phrases which do not specify what has been learned. It would be desirable to know how students trained to identify aircraft actually do visualize them (if they can) and what the characteristics of these images are. Do the drawings show the same features of shape which the students were taught? Are there consistent differences between the drawings of a given airplane and its correct shape? Can such differences be interpreted as exaggerations of certain features and elimination of others? Are the drawings unique and clearly differentiated from one another, as would be predicted in theory? Would the exercise of drawing aircraft be of potential value for training? In order to answer these questions, an experiment was undertaken involving freehand drawings of a number of the aircraft learned in the recognition course.

Assumptions. It was assumed as a basis for this experiment that the drawing of an object from memory is indicative of the memory image itself, i. e., represents the ability to visualize the object. There was no way of checking this hypothesis directly. It was also assumed that the drawings could be scored or rated for the adequacy with which they represented the aircraft. The consistency of this measure, at least, could be checked by comparing the values assigned by independent raters. As a part of this assumption, it was supposed that the "adequacy" of the drawings could be distinguished from the skill or technique which they exhibited. The objection has been raised that freehand drawings will sometimes fail to be adequate representations of the original objects because of a failure in draftsmanship rather than a failure in visualizing. The drawings of this experiment were therefore also rated by independent judges for draftsmanship in order to see with what success this could be done, and in order to compare both the draftsmanship and the adequacy ratings with an external criterion of proficiency in aircraft recognition. One final assumption was made, namely, that the method of constructing a composite drawing from the single drawings of individuals yielded a picture showing the trend of the images of all the individuals.

Method. The individuals used in this study were five classes in aircraft recognition, totaling 196 students, who had just com-

pleted the 30-hour course in Preflight School. The method employed required them to make freehand drawings from memory of 8 airplanes selected from those they had learned. No practice in drawing of any sort had been given them during the course. Three views, plan, passing, and head-on, were drawn of each plane. The 4,700 drawings which resulted constituted the data of this experiment.

Three instructors cooperated in the experiment. From the beginning of the course they had employed in introducing new planes a standard set of phrases and terms to describe the characteristic shape of each. This list of what were believed to be the important features of each plane had been agreed upon beforehand, after discussion in conference with the experimenter. Except for this standardization of descriptive terminology, the course was taught by the usual training methods prescribed. The characteristics or features of these shapes were pointed out merely as cues or tips for accurate recognition. They consisted of such phrases as "beer-bottle fuselage," "oval cowling," "semi-elliptical wing shape," "short nose," "low midwing with delayed dihedral," and other terms descriptive of shape or structure. Terms of this sort had been evolved by instructors in the course of many hours of teaching, but had not been standardized. From six to eight phrases were agreed upon for each plane which were believed to be the significant or important characteristics of that plane. They are listed in table 7.5. It should be especially noted that this standardization of terminology did *not* involve any emphasis on *learning* the features of the planes. The purpose of the experiment was not to test the value of using them in teaching. The features were agreed upon in advance so as to provide a basis for subsequently grading the drawings of the students, and to increase uniformity of teaching in all five of the classes. If there should be evidence that these features affected the drawings ultimately made, it would be because the students had perceived them when they were pointed out and incorporated them into their visual images spontaneously. The major part of the work of the course consisted of the repeated presentation of projected slides and their identification by the "review" technique.

At the end of the 30 hours of instruction each student was given a booklet of eight sheets of paper on which to draw three views of each of 8 planes. A single sheet was divided horizontally into three parts labelled "Plan View," "Passing View," and "Head-on View." The size of each drawing was kept constant by marks which fixed the length or wingspan, but not the proportion or shape. All drawings faced the same way. These requirements were necessary in order to make composites. The 8 planes were selected from the 40 taught with consideration for their familiar-

Table 7.5.—*The significant visual features of the aircraft's shapes and their inclusion in the drawings of students*

Plane	*Significant features*	*Inclusion or absence of feature in composite drawing*	*Percent of individual drawings which include the feature (N = 196)*
B-26	Tapered wings	Present	48
	Long nose and nacelles	Absent	39
	Tapered tailplane	Present	51
	Cigar fuselage	Present	77
	High rounded vertical	Present	78
	Round fuselage and nacelles	Emphasized	89
	Straight wing	Present	68
	Dihedral of tailplane	Emphasized	64
Spitfire	Elliptical wing	Present	87
	Round tailplane	Absent	37
	Flat arrow fuselage	Emphasized	71
	Rounded vertical	Emphasized	82
	Low wing with dihedral	Present	42
	Rectangular radiators below wings	Present	40
Zeke	Broad chord, even taper	Present	52
	Tapered tailplane	Absent	39
	Bubble canopy	Emphasized	85
	Triangular vertical	Present	50
	Flat belly line	Present	78
	Round cowling	Emphasized	95
	Low wing with dihedral	Present	57
He-111	Nose and engines lined up	Absent	12
	Notched wings	Emphasized	98
	Elliptical tailplane	Absent	26
	Elliptical vertical	Present	72
	Round back	Present	61
	Bathtub in belly	Emphasized	73
	In-line engines	Absent	18
	Dihedral of wing	Present	33
P-47	Semi-elliptical wing	Present	43
	Birdtail stabilizer	Absent	18
	Beer-bottle fuselage	Emphasized	59
	Half-diamond tail	Present	66
	Round belly line	Present	49
	Low midwing	Present	70
	Oval cowling	Emphasized	86
Typhoon	Short nose	Present	39
	Even-taper, round tip wings	Present	42
	Stabilizer shaped like wing	Absent	18
	Bullet nose	Present	74
	Large airscoop	Emphasized	95
	Bubble canopy	Emphasized	48
	Round vertical	Emphasized	69
	Delayed dihedral	Absent	41
P-51	Even-tapered wing	Present	86
	Clipped wing and stabilizer	Present	95
	Clipped vertical	Present	95
	Bullet nose	Present	75
	Deep airscoop	Emphasized	74
	Flat, oval cowling	Present	62
	Low wing with dihedral	Present	50
FW-190	Short radial nose	Present	57
	Square tip wing, even taper	Present	52
	Rectangular stabilizer	Present	72
	Flat belly line	Present	64
	Diamond cockpit	Absent	40
	Round vertical	Present	57
	Round fuselage	Emphasized	91
	Low dihedral wing	Present	49

ity and with regard to their being adequately characterized by the set of named features ascribed to them. They were the B-26, Spitfire, Zeke, He-111, P-47, Typhoon, P-51, and FW-190.

Rating of Drawings for Adequacy. The drawings were analyzed in two ways, first by rating the adequacy with which they represented the characteristics of the shape of the real plane, and second by constructing composite drawings for each view of each

plane. A semi-objective method of obtaining the "adequacy" rating was adopted. This consisted of counting the number of characterizing features which were shown on the drawings, taking all three views together. The rater looked at silhouettes of the real planes while deciding whether a given characteristic was present or absent in the drawings. This score for each plane was totaled for all 8 planes and the resulting number was called an "adequacy" score for each student. If the terms agreed upon by the instructors can be assumed actually to characterize the shapes, this score is in fact a measure of the adequacy of the drawings. The objectivity of this method of rating is indicated by a reliability of .78 between the scores assigned by two raters, working independently, on the drawings of the B-26 by 100 subjects.

In addition to this rating a separate rating was made on a 10-point scale, by each of three judges, of the draftsmanship exhibited in each set of drawings, irrespective of the correctness of the shapes. The rating was based on whether the drawing was a skillful reproduction of a plane in the abstract rather than of the particular plane it was intended to represent. The raters did not look at real silhouettes and were unfamiliar with the real planes. The reliability of these ratings is indicated by correlations of .89, .87, and .85 between the ratings of the three judges, on all 196 students. A rating on draftsmanship was then obtained for each student.

Relation Between Ability to Draw the Airplanes and Ability to Identify Them. In addition to the "adequacy" score described above, the final course grades were obtained for each student. This grade was a measure of the ability to identify aircraft from slide photographs as determined by the final slide examination and other tests given previously. For administrative reasons, unfortunately, the motion picture proficiency test could not be administered to this group of students.

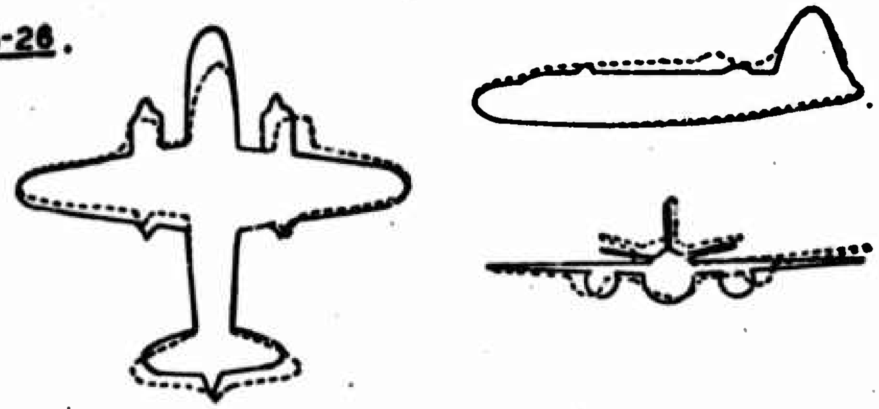
The product moment correlation between the "adequacy" score and the final grade for the group of 196 students was found to be .61. The reliability of the "adequacy" score was found to be .78, between two independent raters. The reliability of the final grades in the aircraft recognition course at the time of this experiment was estimated to be in the neighborhood of .85, from previous computations. Taking reliability into account, the coefficient of .61 between such superficially disparate achievements is striking. It is a correlation between ability to visualize aircraft and ability to identify them overtly. As such, it has interesting theoretical implications. It implies that there is a fairly close connection between the development of differentiated experiences or images of airplanes and the development of discriminative responses to them.

The objection is sometimes raised to the method of requiring free-hand drawings from unskilled persons that they cannot be judged for adequacy of representation because they will be so largely reflections merely of drawing technique or "artistic ability." The raters in this experiment did not find it difficult to judge "adequacy" in terms of a specific set of characteristics or features of the original shape. They also did not find it difficult to make ratings of draftsmanship. But in order to obtain further evidence on this question, the "adequacy" ratings were correlated with the draftsmanship ratings and their correlations with the final grades were compared. The relationship between the two ratings is .54, but whereas the former correlates .61 with grades the latter is correlated with them only to the extent of .37. Moreover, this relationship of draftsmanship to grades drops to insignificance (.07) when the influence of the "adequacy" score is partialled out. It may be concluded that, although draftsmanship is related to the adequacy of the drawings, as indeed it would be expected to be, the two may reasonably be distinguished, and that only the latter is related to recognition proficiency independently of the other.

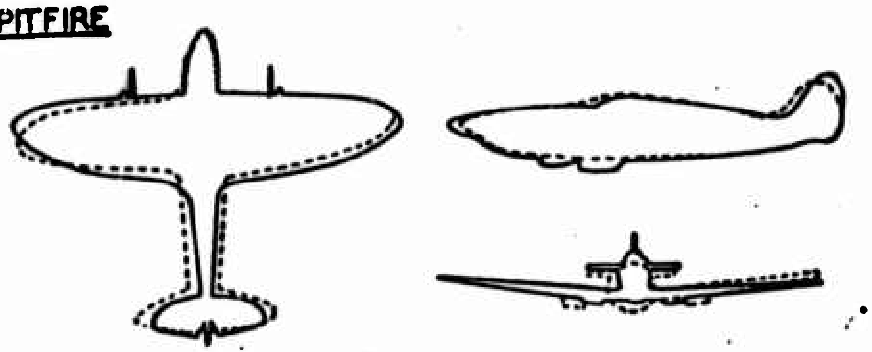
The Composite Drawings. Composite drawings were constructed by a technique of tracing the original drawings on translucent paper, five at a time, and then averaging the five lines by eye on another piece of translucent paper. It was discovered, and repeatedly confirmed, that a sample of as few as 25 papers was sufficient to yield a composite which was representative of the full population of 196 papers. Composites obtained from groups of 25 papers were practically identical with one another. Two individuals performed this visual tracing and averaging and the method was checked frequently by comparing the composites obtained by the two individuals working separately. Agreement was very high. There resulted from this procedure three views of each of the 8 airplanes, or 24 composites. These outlines were then superposed over the outlines of the real airplane silhouettes, using a photographic enlarger to bring one outline to the same size as the other. The shape of the composite could then be readily compared with the shape of the real object. The superposed composites are shown in figure 7.1. The correct outlines are shown in solid lines and the composite drawings in dotted lines. Table 7.5 gives for each plane the list of terms which characterize it, and whether each feature is present, absent, or emphasized in the composite drawings. Also listed for each feature is the *percent* of the individual drawings in which it was counted as present.

Differentiation of Memory Images. An inspection of the drawings of the eight planes, and especially a comparison of the different composites, made it evident that the students had learned to

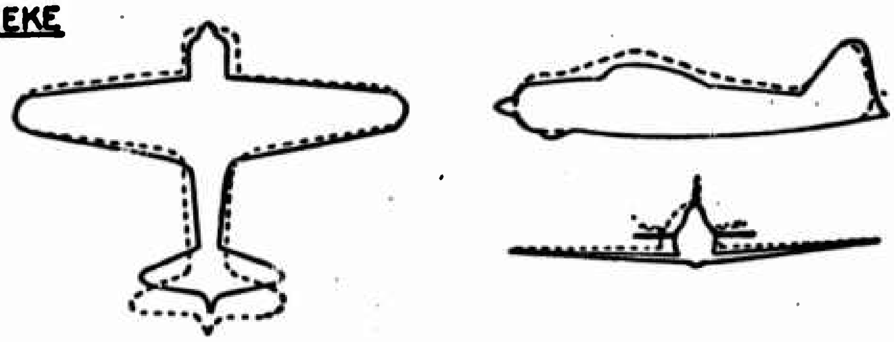
B-26.



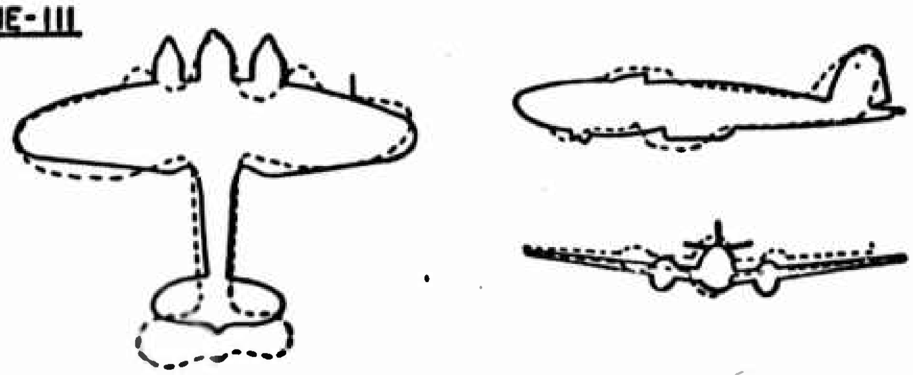
SPITFIRE

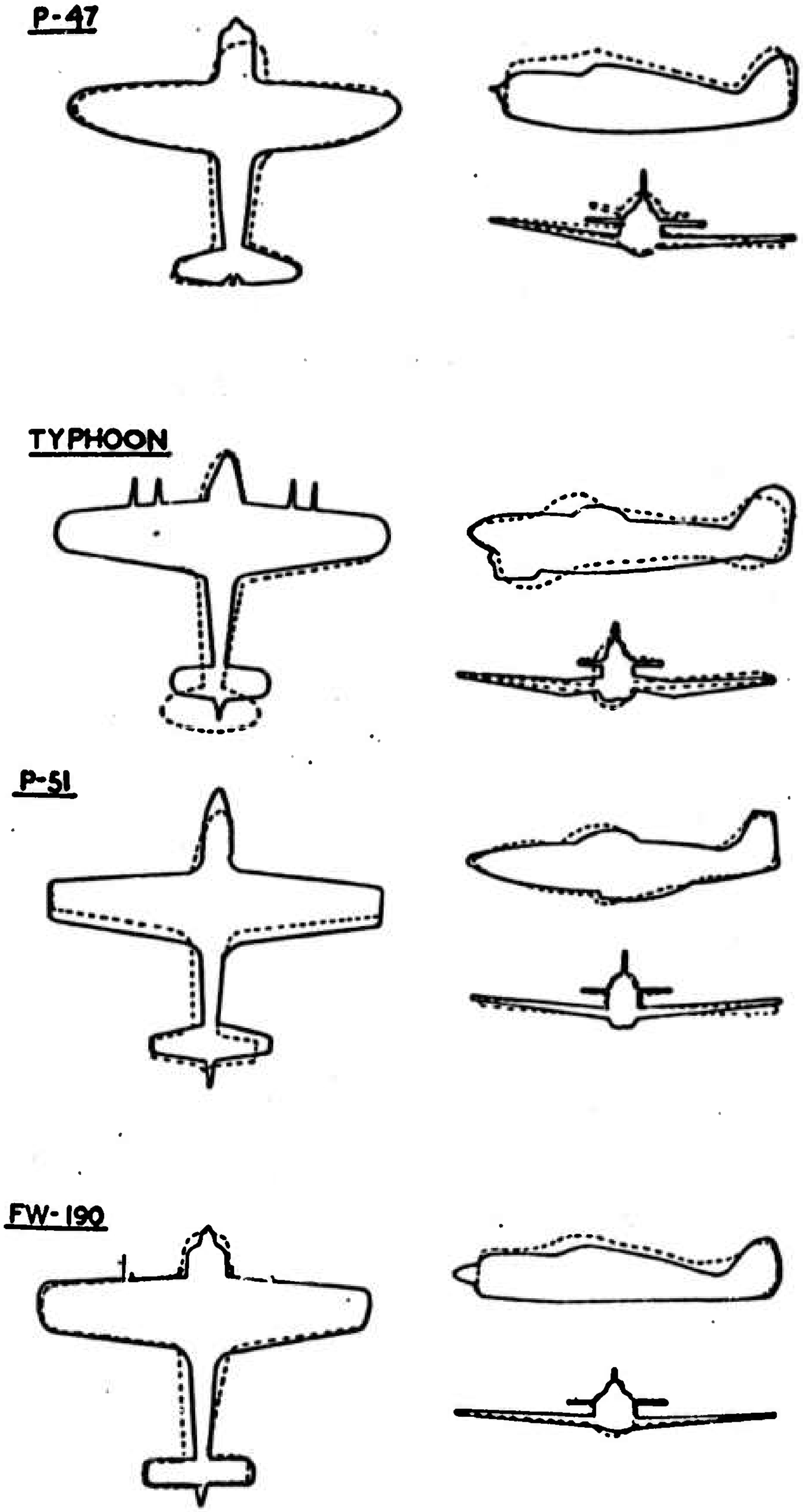


ZEKE



HE-111





FIGURES 7.1A and 7.1B.—Composite remembered shapes as compared with real shapes.

visualize the airplanes as unique and clearly differentiated from one another. The composites, and the great majority of the original drawings, had the main visual characteristics of the planes they represented, although in many cases they were crude and unskillful as drawings. There was no tendency for the composites to degenerate into representations of a generalized airplane or to "converge" toward a single type. They were in fact specific. It may be inferred that differential visual images of the airplanes had been acquired by the students.

Differences Between Drawings and Planes. There were, however, as evidenced by the composites, consistent differences between the drawing of a plane and its actual shape. The composites were by no means copies of the silhouettes. This was true, in greater or lesser degree, for all views of all planes. There are clear indications in the composites of "constant errors" in the typical visualizations of each aircraft, but in general they were not errors of a sort that would lead to mis-identification of the plane. The fact to be especially noted is that the errors of the individual students, when averaged in the composites, did not cancel one another out but showed a consistent trend.

Exaggeration of Characteristic Features. The differences between the composites and the real shapes suggest a number of hypotheses which are significant for the theory of perception and memory but are not relevant to this report. The fact of importance to aircraft recognition learning is that in many instances the differences can be accounted for as *enlargements or exaggerations of those features of the shapes believed by the instructors to be characteristic and described by them in class.* Some of the composites are even suggestive of caricatures of the planes. If these characteristics have modified the images in these ways despite the contrary influence of the stimulus patterns, then it may be supposed that these features have played an important part in differentiating the various memory-images from one another and in giving them significance or meaning. The importance of these descriptive features for learning to visualize, and presumably also to identify, airplanes is clearly implied. Since it is also true that the "organization" of the images involves or is determined by these features, the fallacy of advocating that aircraft be taught only as "total forms" without reference to their unique, specific, and determining characteristics is evident.

The use of caricatures of planes for instructional purposes in aircraft recognition had often been suggested and had been adopted by some schools and stations. A number of them were examined in connection with the present study but proved to be disappointing as caricatures. The results described suggested that some features are more amenable than other to pictorial emphasis,

that some are more dominant than others in the mental pictures of students, and that those features which appeared in emphasized form in the composite drawings were ones which could most effectively be emphasized in caricatures and posters.

The Potential Value of Drawing Exercises. In view of the evidence presented it was suggested that training in the drawing from memory of the airplanes to be learned in a recognition course would be a useful method of training supplementary to the usual practice in identifying the airplanes from photographic slides or other presentations. Since it seemed likely that the shapes of aircraft become identifiable by a process which has not only the aspect of learning discriminative responses but also the aspect of acquiring differentiated images, the use of methods appropriate to this latter aspect of perceptual learning should not be overlooked.

The Effectiveness of Differential Reinforcement in Learning to Identify Aircraft

Introduction. Emphasis on the flash method of training in AAF schools and in particular on the progressive speeding up of the exposure interval had been reduced by the early part of 1944 by direction of Headquarters, AAF, in Washington. Supplementary devices for training were coming into increasing use, such as training films, photographs, models and "shadowgraph" presentation of models. It seemed likely that, although flash training as such was ineffective, there was one feature of training with film-slide projectors and slide-kits which was highly effective, namely the practice of daily slide reviews or tests. This practice had the virtues of frequent repetition, knowledge of the progress of one's learning, competitive attitudes among the students, and above all the advantage of providing immediate confirmation of right responses and correction of wrong responses. Persistent errors and confusions between airplanes which were difficult to discriminate could in this way, and only in this way, be eliminated.

Experiments in the psychological laboratory had shown that the discrimination of mutually similar stimulus objects could be achieved through *differential reinforcement* of right and wrong responses to these objects. In theory, reinforcement was positive in the case of a confirmed response and negative in the case of an unconfirmed response. The final result of the learning was a set of differential responses. It was the hypothesis of the Film Unit that learning to identify aircraft by shape was, in one of its aspects, a process of this sort.

The practice of immediate confirmation or correction of responses could be contrasted with a practice which consisted of presenting the stimulus objects and their responses (names) for

association. It might be assumed that the learning of visual shapes occurred simply by a process of *perceiving* them repeatedly. If this were true, it would be sufficient to look at airplanes and think of their names, or to present the learner with a repeated series of visual impressions in connection with the names. This latter assumption underlay the so-called WEFT system of study, and was also characteristic of training films on aircraft recognition. The procedure was also very easy to adopt in using photographs or models for training. In order to verify the hypothesis that it was ineffective in comparison to the former practice, an experiment was performed. The procedure of providing immediate differential reinforcement of responses was compared with a procedure in which the slide-pictures were presented *for the same amount of time* but without confirmation or correction of responses.

Method. The individuals used in this experiment were 280 aviation trainees in the Classification Center, Santa Ana Army Air Base, prior to their entry into the Preflight School. They were divided for testing purposes into sections of approximately thirty men each. All of these sections learned 20 slides of 20 unfamiliar foreign aircraft, half of them learning by one method of presentation and half by the other method. Since the experiment was designed to measure amount of learning rather than recognition proficiency, those individuals who were found in advance of the experiment to be familiar with as many as two of the foreign planes shown were eliminated from the experiment.

The twenty slides, representing various foreign planes, were selected on the basis of their unfamiliarity. They were largely obscure, nonoperational, foreign types. An attempt was made to select slides that had no distinguishing background features, so as to prevent the recognition of slides rather than planes. The planes represented had, with two exceptions, names (*e. g., Arado, Fulmar, Cant Z*) rather than numbered designations.

Both groups of trainees were required to learn the names of the 20 planes during three presentations in a prearranged random order. The fourth presentation of the slides served as a test of the amount of learning which had taken place. In detail, the training methods employed were as follows:

Visual impression without reinforced responses. During each of the three presentations, each of the 20 slides was shown for five seconds. The name of each plane was announced just before it appeared on the screen and was repeated while the name was on the screen. On the fourth presentation, each slide was exposed for only 2.5 seconds, and the trainees were instructed to write down the name of each plane on appropriately numbered answer

sheets. Guessing was encouraged. The score on the test was simply the number of planes correctly identified.

Visual impression with reinforced responses. The first time the 20 slides were shown, each was exposed for 5 seconds and the name of each plane was announced before and during the exposure. On the second and third presentations, however, the slides were exposed for 2.5 seconds, after which the trainees were required to identify each in writing on a numbered answer sheet. Guessing was encouraged. In case no identification could be made, a line had to be drawn through the appropriate answer space. Immediately thereafter the same slide was exposed again for 2.5 seconds and its name was announced. Individuals who had made a correct response placed a check mark beside it; those who had made either an incorrect response or none at all were required to write down the correct name in the appropriate blank. The fourth presentation, constituting the test, was conducted in a manner exactly similar to that described above. It should be noted that the *intention to learn* was equally strong in the two groups.

The plan of the experiment is shown in the following diagram:

Trial One	*Trial Two*	*Trial Three*	*Trial Four*
Non-reinforced: a. Name b. Plane (5 sec.) with name	a. Name b. Plane (5 sec.) with name	a. Name b. Plane (5 sec.) with name	a. Plane (2.5 sec.). b. Written response.
Reinforced: a. Name b. Plane (5 sec.) with name	a. Plane (2.5 sec.) b. Response attempt c. Plane (2.5 sec.) with name d. Completing, confirming, or correcting of response	a. Plane (2.5 sec.) b. Response attempt c. Plane (2.5 sec.) with name d. Completing, confirming or correcting of response	a. Plane (2.5 sec.). b. Written response.

Results. The average number of planes correctly identified by each group, together with its standard deviation, is given in table 7.6. After three trials, the group which learned by the unrein-

TABLE 7.6.—*Amount of learning by two methods*

	Number of individuals	*Average score*	*SD*	*Percent of individuals getting all slides correct*
Unreinforced method...............	149	10.49	4.70	2.7
Reinforced method.................	131	14.56	4.15	13.7

forced method identified an average of 10.49 planes out of 20 correctly, while the group which learned by the reinforced method averaged 14.56. The latter group, therefore, reached a proficiency which was more than 40 percent superior to that of the former group. The critical ratio of the difference between these averages is 7.69, indicating that the difference is highly significant. Only a small proportion of each group learned all 20 of the planes in the short training period given. But the last column of figures shows that the number of trainees in the reinforced group who

learned all the planes was over 5 times as great as the number who accomplished this in the unreinforced group.

Discussion. The experiment reported showed clearly that the principle of *overt response with reinforcement*, the effectiveness of which had been demonstrated many times in learning studies, was applicable to the special form of visual learning required for the successful recognizing of aircraft. There was little novelty in this principle as a commonsense rule for teaching and training. It implied that active recitation or reaction was a good thing when students were required to check their answers, since it allowed them to perceive their mistakes while the latter were still clearly in mind, and also to confirm the successful answers immediately. The danger was that this principle would not be fully utilized in a new field of training and teaching—a field in which methods of instruction were being vigorously debated and rejected.

The results of the experiment were interpreted as a demonstration of the superior effectiveness of any classroom procedure which required active identifying of responses and permitted the confirming or correcting of these responses. The "Flash System of Instant Identification" had in actual practice made use of this procedure without acknowledging its importance. The studying of aircraft by the so-called WEFT system had, as a rule, not done so. The practice of looking at aircraft shapes or visually impressing them on the learner was not enough; learning was more effective if differential responses were required. The implication was that the showing of models, motion pictures, posters, "planes-of-the-day," and in fact any other type of material would be of little value if they were presented merely as *exhibitions*. The great volume of visual training aids being produced could not be expected to yield the most effective result unless they were employed as positive instruments of instruction rather than for display.

The application of the rule to posters of aircraft suggested that they be designed without the name of the plane being prominently shown but with the caption, "Do You Know This Plane?" and with the answer printed in small type at the bottom or under a small flap which would have to be raised. The application to the display of models suggested the plan of suspending them from the ceiling on numbered hooks which could be changed from day to day and requiring students to identify the airplanes by successive numbers.

The application to the showing of training films on aircraft recognition suggested an even more radical departure from common practice. Motion pictures were, and are, primarily exhibitions or displays. The learner remains passive during the showing of an ordinary film and, however vivid his visual impressions, cannot practice what he is learning. The motion picture pro-

ficiency tests already produced required active responses but did not permit immediate reinforcement of the responses. They were therefore not suitable for reviews. The film could, of course, be shown a second time after the test proper and the students could then correct their own answers. The most effective type of training film, however, would be one constructed in the following sequence: item number, view of the airplane to be identified, interval for response, another view of the airplane (possibly a "stop-frame") with announcement of its name and possibly mention of a visible identifying feature, during which interval the preceding response is checked or marked wrong; second item number, view of another airplane, etc. Such a film would be similar to slide reviews except that the planes would be presented in motion and in changing attitudes. It could be stopped at any point for discussion or for the answering of questions. The semi-illuminated classrooms utilized in recognition teaching would be adequate for showing it. The Film Unit therefore recommended that this type of training film be produced. In view of its complexity and unconventional form, however, only the type constituting a test proper (incorporating only the first three elements of the sequence recommended) was actually produced for review purposes in the classroom. They were called "Testcraft" films.

Conclusions. Learning to identify aircraft by shape is markedly facilitated by differential reinforcement of right and wrong naming responses. Methods of training which rest on the assumption that such learning is a matter of visual impression and the spontaneous acquiring of memory images without any necessity for overt reinforced responses are, by implication, relatively ineffective for this kind of performance.

The Relative Effectiveness of Teaching Similar or Dissimilar Planes Together

Introduction. The organization of the course in aircraft recognition was a matter of considerable importance to those whose responsibility it was to plan the most economical use of classroom time. Since the material to be learned consisted of forty planes, in some respects much alike and in some respects different, one of the problems of organization was that of the order in which the planes were to be learned. Four nationalities, American, British, German, and Japanese, and three main classes, four-engine, twin-engine, and single-engine, were represented. There were three major alternative orders of presentation for this group of forty planes:

Presenting Dissimilar Planes Together. The practice in preflight school was to introduce two dissimilar planes of the same nationality each day. Thus, the F6F, a single-engine fighter, was

703313 47--11

presented along with a twin-motored bomber, the B-25; the B-34 was introduced during the same class hour as the P-47, etc. The justification made for this procedure was that it prevented confusion. It did not, of course, do so permanently. The effect of this procedure was only to minimize confusion in the *early stage of learning the list of airplanes*, or in other words, to defer it.

Presenting Similar Planes Together. The minimizing of confusion between similar planes was a very desirable goal for recognition training. In a previous experiment it was pointed out that one of the most commonly-expressed questions by students was "How can I tell one airplane from another which looks like it?" Confusable airplanes, although not of the same nationality, practically always belonged to the same general *class* of airplanes (four, two, or one-engined). Some instructors believed that the best way to prevent such confusion was to present similar planes in pairs, regardless of nationality, and to point out clearly those features of the aircraft by which they could be distinguished from one another. In practice this would be accomplished by introducing pairs of similar airplanes during the same class period. Whether emphasis on the differences of these similar pairs would be the more efficient method of teaching recognition, or whether it would interfere with the learning of the total list of forty planes, was a question which could be answered by an experimental study.

Presenting Planes According to a Systematic Classification. It would have been theoretically possible to arrange the planes to be taught in a systematic order which emphasized both similarities and differences. This method would require that planes be divided first into a number of classes on the basis of common features. Within each of these classes the discriminating characteristics of the planes could be emphasized by division into subclasses. For example, in the class of four-engine planes, a certain subclass would contain planes which are similar in that they have twin tails, while another subclass would have single tails in common. The planes in each of these subclasses would be dissimilar in some respects (e. g., shape of verticals), which would serve to distinguish one from another. This type of presentation will be taken up in the next section. It could not, however, be incorporated into the present experiment for various reasons, chief of which was a resistance by the authorities to introducing more than two new airplanes in each class period.

The experiment to be reported was designed to show which of two sequences of the airplanes constituting a course in aircraft recognition was the more efficient for training. In one sequence, two planes of the same nationality which were very *dissimilar* were introduced during each class period; in the other sequence,

two planes which were *confusable*, irrespective of nationality, were presented during each class period.

Method. Six classes in aircraft recognition at the AAF Preflight School (Pilot), Santa Ana Army Air Base, were employed in the experiment. These classes, containing about 30 men each, were divided into two groups which were matched for initial knowledge of airplanes on the basis of scores on a pretest of 30 slides of American aircraft. Three instructors taught these classes, each instructor having one class in each group, so that differences in teaching ability could not affect the results.

Both groups of classes were taught the same forty planes by the usual method. Equal numbers of slides and equal amounts of time were used in the reviews for each group. Slides were reviewed at an interval of one-tenth second. One group of classes was taught the planes in an order which required the presentation of two dissimilar planes of the same nationality each day, (e. g., P-47 and B-34, Spitfire and Mosquito, etc.). American planes were taught first, then German, then Japanese, and finally British. In the second group, two similar and confusable planes were presented during each class hour (e. g., P-51 and Me109, Spitfire and Hurricane, etc.), without regard to their nationalities. The training period for both of these groups was 26 hours. At the end of this period both groups of classes were tested by means of a slide examination composed of 45 slides showing the 40 planes presented during the course. In addition, they were tested with the Aircraft Recognition Proficiency Test (Preflight Level). The pretest on which the groups were matched correlated 0.52 with the final slide examination and 0.62 with the motion picture test.

It should be pointed out that it was impossible to prevent the students who were taught planes in the "dissimilar" order from studying recognition manuals outside of class, and thus making comparisons between confusable pairs. However, neither could the group to whom similar planes were presented be prevented from studying them outside of class in some grouping different from that given in class.

Results. The average number of planes correctly identified out of 45 on the final slide examination was 36.24 ($SD = 3.23$; $N = 93$) for the group to which similar planes had been taught. For the group which had been presented dissimilar planes together, the average score on the same examination was 35.79 ($SD = 3.84$; $N = 90$). The critical ration of the difference between these two averages, corrected to allow for the reduction of differences which results from using matched groups, is 1.01. This value indicates that a difference of this size, in favor of either group, could be expected to occur by chance as many as 31 times out of 100. Consequently, the difference which appears is not a significant one.

The group to which similar planes had been taught together obtained an average score of 63.09 ($SD = 12.00$; $N = 92$) on the motion picture test, which was assumed to be the more trustworthy measure of proficiency. The average score of the second group, to which dissimilar planes had been presented together during each class period, was 61.38 ($SD = 12.16$; $N = 80$). Allowing for the reduction of differences produced by matching the two groups, the difference between these average scores has a critical ratio of 1.17. A difference of this size might occur by chance 24 times out of every 100 similar experiments. In the case of this criterion of proficiency also, it may be concluded that no real difference had been demonstrated between the groups as a result of their differential training. However, it was suggestive for future research that the slight advantage obtained with both tests was in favor of the group for which similar planes were presented together.

Conclusions. The prevalent belief that confusion should be minimized by separating the similar aircraft in time and presenting dissimilar pairs together was not confirmed by the experiment.

The opposite hypothesis that confusion could be minimized by pairing and contrasting similar aircraft was also not confirmed, except by a slight trend in the results.

A more fundamental thesis, to which the latter was only a partial approach, would predict that confusion could best be minimized by organizing aircraft in classes and subclasses of visual similarity. This hypothesis was not tested in the experiment. Since the mere pairing of similar aircraft (leaving many others unpaired which are also similar) is hardly a test of the value of organizing and classifying aircraft by the visual features which are similar and those which are unique, the experiment suggests that had the latter procedure been tried out the results might have been positive.

The Organization of the Aircraft Recognition Course

The experiment last reported, together with the evidence from the experiment on "total form" *vs.* "features," suggested that the persistent difficulty of discriminating between similar shapes might be met by a reorganization of the aircraft recognition course. Several steps would be necessary for such a reorganization. First, an analysis should be made of the *identifying features* of the list of aircraft to be learned—not simply of the *descriptive* features which had characterized the WEFT analysis. The characteristics which were distinguishing should be isolated. Second, these distinguishing features should be used to make a classification of the aircraft by shapes, such as would permit a conceptual organization of their similarities and differences. Conceptualiza-

tion of material should promote differential learning. Third, the classes and subclasses of aircraft should be made the basis of organization of an experimental course, all the aircraft being systematically presented in the early part of the course (at a much faster rate than two per day) and the remainder of the course being devoted to differentiating and reviewing them.

The first two steps of this project were carried out. A beginning had been made on the problem of isolating the distinguishing features of aircraft which were most significant for learning during the "features" experiment and the "drawing" experiment. A group of instructors in recognition had conferred with the experimenter and had come to an agreement on a standard terminology of descriptive features. Those which were identifying were then analyzed by the experimenter in the course of constructing a large chart. The chart was constructed in the following way.

As a start, it could be stated with some degree of assurance that the members of a group of four-engined planes are more similar to each other than they are to any other planes of the total list. Among this group of four-engined planes some members are more apt to be confused than others. The B-17 is more like the C-54, which has a single tail, than it is like the B-24, which has twin tails. Thus two subgroups may be distinguished within the total group. The features of each member of the subgroup which distinguish it from each other member can then be clearly pointed out. The remaining groups and subgroups to be outlined were arrived at in a similar manner. It would of course have been desirable to have had exact quantitative measures of the degree of confusability (generalization) between the planes, but these were not considered essential. The resulting groups were divided into seven, in order to make the number of planes in each group small enough to be capable of presentation within an ordinary class period. These groups were as follows, with the distinguishing features underlined for each plane:

FOUR ENGINES
(LESSON 1)

I. Single Tail—
- B-17: *Large Faired Vertical*, Round Fuselage
- C-54: *Small Narrow Vertical*, Long Nose

II. Twin Tails—
- Lancaster: *Pear-shaped Vertical*, low-horizontal
- Halifax: *Triangular Vertical*, high-horizontal
- B-24: *Oval Vertical*, long narrow wing

TWO ENGINES
(LESSON 2)

I. Twin Tails—
- *Squared Verticals*
  - B-25: *Modified Gull Wing*, slab-sided nose
  - Do217: *High Straight Wing*, deep bulbous nose

Egg Shaped Verticals

B-34: *Deep fuselage*, pointed wing tips

Me110: *Slim fuselage with long cockpit*, square wing tips

P-38: *Twin booms*

(LESSON 3)

II. Single Tail—

A. *Long Nose—*

1. Slab-sided Fuselage

A-20: *High Sloping vertical*

Wellington: *Narrow triangular vertical*

2. Cylindrical Fuselage

C-46: *Curved nose without step*, rounded vertical

C-47: *One step in nose*, straight-edged vertical

B-26: *One step aft of nose*, U-shaped vertical

Sally: *Two cockpits*, narrow angular vertical

Betty: *Short blister cockpit*, wide triangular vertical

(LESSON 4)

B. *Short Nose—*

1. Nacelles Even with Nose

He—111: *Elliptical vertical and horizontal*, notched wing roots

Ju-88: *Vertical beyond tail*, bulbous cockpit

Mosquito: *High-pointed vertical*, long curved horizontal

2. Nacelles Project Ahead

Beaufighter: *Triangular Vertical*, straight top with blister

Me-210: *Tall "U" Vertical*, high arched cockpit

SINGLE ENGINE

(LESSON 5)

I. Pointed Nose—

A. *Long Nose—*

1. Square-Top Vertical

Ju87: *Inverted Gull Wing*, fixed landing gear

P-51: *Square Wing Tips*, shallow belly scoop

2. Round-Top Vertical

P-39: *Clean Belly Line*, blister cockpit

P-40: *Deep nose scoop*, straight leading edge

(LESSON 6)

B. *Short Nose—*

1. Blister Cockpit

Typhoon: *Large scoop under nose*, dihedral in outer wing

2. Humpback Cockpit

Hurricane: *Scoop under cockpit*, rounded fin and rudder

3. Streamlined Cockpit

Spitfire: *Underslung nose*, elliptical wings

Me-109: *Thick bullet nose*, high horizontal

Tony: *Thin round nose*, scoop under cockpit

II. Blunt Nose—

A. *Mid-Wing*—

1. Square Wing-Tips

TBF: *Long Cockpit,* prominent belly step

F6F: *Short Cockpit. rounded horizontal,* oval fuselage

F4F: *Short Cockpit, square-tip horizontal,* round fuselage

2. Semi-Elliptical Wing

P–47: *Triangular Vertical,* oval fuselage

B. *Inverted Gull Wing*—

F4U: *Inverted "U" Vertical*

C. *Low Wing*—

1. Tunnel Cockpit

Val: Elliptical wing, fixed landing gear

2. Blister Cockpit

Oscar: *Straight leading edge,* faired vertical

Zeke: *Tapered leading edge,* triangular round-top vertical

3. Low Angular Cockpit

FW–190: *Narrow rectangular horizontal,* short nose

This outline is not to be construed as a "system" for the identification of airplanes. It could not practically be used for that purpose, since under actual conditions of recognition, there was no guarantee that the particular distinguishing feature chosen would be evident (e. g., a rectangular horizontal cannot be seen in a head-on view). Neither should it be objected that a student who has learned this particular organization would have to stop and run through it verbally in order to recognize a particular plane which he sees in the sky. It is obvious that only a good deal of reinforced practice of the specific discrimination habits themselves would lead eventually to rapid and accurate recognition. No verbal system could accomplish this by itself; the course had to be taught by a method which made the student see and recognize as many views of each plane as large a number of times as was practically possible.

The outline did represent a proposed organization for the presentation of planes, based upon an analysis of similarities and differences between them. It was conceived to be an aid to the learning of a set of planes, particularly in the *initial stages* of training. If discriminations between the planes to be taught were established more rapidly by means of verbal rules such as these, there was reason to believe that training would be more efficient, and that a greater degree of proficiency would be the result.

This proposal to emphasize in training a number of distinguishing features of aircraft suggested to some instructors the WEFT system. It was pointed out that the features given in the outline

were not intended to replace practice in recognition of aircraft, but merely to act as aids in the initial establishment of discriminations between similar planes. These features were about the minimum number which could serve to distinguish confusable plane shapes. They were to be contrasted with the features of the WEFT system, most of which did not perform this function and could on this account be said to waste time. If utilized at the beginning of training in the manner proposed, the present set of features would, it was suggested, permit *more* time to be spent in practicing recognition of "total forms" of planes.

An experiment to compare the proposed type of course with the regular course was not possible to arrange, for administrative reasons. Instead, the experimenter had to be content with a tryout of the method in a small informal course given to a mixed group of subjects ignorant of aircraft recognition, acting himself as the instructor. Although no statistical comparisons were possible, the procedure appeared to be practicable and in fact very successful. The method unfortunately could not be pursued further because of the pressure of other research and the termination of the entire project when training in the AAF was sharply cut.

EXPERIMENTAL STUDIES CONNECTED WITH THE PERCEPTION OF AIRCRAFT AT A DISTANCE

The fact that aircraft in combat normally had to be identified at distances upward of half a mile was one of the difficulties in the training program. Instruction had to be carried out in a *miniature situation*, using pictures, projected images, or models. The miniature situation should, if possible, reproduce the impression of an airplane at a distance, but this requirement presented a problem in the psychology of distance perception. The problem was complicated by the fact that estimating the range of an airplane seen against a background of sky is a vague and difficult matter at best —a fact which is discussed in greater detail in Chapter 9. The size of an airplane likewise could not be accurately represented in a a picture having only sky or clouds as a background. Perception of size under natural conditions in the air is as difficult as perception of distance; on this account the sizes, or wingspans, of aircraft had to be memorized by students. The problem of the *portrayal* of distance in relation to size, so far as it was possible at all, was a problem in need of experimental study.

Another question in need of evidence was at what distances aircraft could be identified. How far away could the shapes be discriminated? The maximum range for effective fire was known, but opinions varied as to the maximum range at which visual identification was possible. Experimental evidence would be difficult to

obtain but even an approximate answer would be useful both practically and theoretically. The experiments to be reported in this section were devoted primarily to these two questions.

Factors Determining the Perceived Range of Airplanes Shown in Projected Slides

Introduction. A number of methods had been proposed for increasing the realism of training in the aircraft recognition course with respect to the impression of distance or range. The slides available at the beginning were very largely photographic "close-ups" of the various aircraft. Since identifying backgrounds had to be eliminated, they eventually came to have no background at all except clear film representing sky or clouds. The first of these methods was the progressive decreasing of the size of the picture flashed on the projection screen—a matter easily accomplished by moving the projector closer to the screen. This practice was at first recommended but later generally abandoned in AAF schools. The fact was that in this situation the students did not see the airplane in the picture as more distant, but saw merely a smaller picture. A second method was that of producing the slides so as to decrease the size of the airplane itself within the picture, i. e., in relation to the surrounding clear area representing sky or clouds. A new series of "distant view" slides were made by a photographic reduction process and distributed by the AAF Training Aids Division which were intended to represent aircraft seen at 1,000 yards. Aircraft were represented at this distance in the sense that they subtended the same visual angle as would a real plane at 1,000 yards. These slides were generally considered superior to the "near-view" slides previously available. A third method was to increase the distance at which students viewed the screen, or to mark out on the floor of the classroom the seating distances which would correspond to the theoretically greater apparent ranges of the aircraft pictured on the screen. A fourth method was the proposal to use projected stereoscopic images in order to give an impression of depth to the picture.

There were several pitfalls in the way of using these methods for their intended purpose. In the first place it was important to know how realistically the 1,000-yard series of slides when projected in the classroom did represent or simulate the range of real aircraft. It was specified in the original directions accompanying them that these 1,000-yard slides would represent the appearance of planes at 1,000 yards only if the rectangular projected image were 28 inches in width (the standard condition) and if students were seated at a fixed distance from the screen. It was a question whether these specifications were necessary. The third method, like the first, and like the specifications for using the distant-view

slides, rested on the assumption that the perceived distance of an object is dependent on the size of the retinal image in the eye of the observer, defined in terms of the visual angle cf intercept. This assumption was not in accordance with the facts of perceptual size constancy (cf. chapter 9). Although it is true that as an object moves away from the observer its retinal size diminishes proportionally and its perceived distance increases, it does not follow that as the retinal size of a picture diminishes, for whatever cause, the represented distance of an object in the picture increases. Ordinary photographic experience indicates that a picture may be enlarged or reduced and held close to or far from the eyes without an appreciable effect on the distance simulated in the picture.

In order to clarify these conceptions, an experiment was performed to discover whether in fact the size of the airplane in the screen picture, or the size of the retinal image as it diminishes with viewing distance, determines apparent range, or whether both do. Another purpose was to discover how determinate the appearance of range in a picture actually is, or at least how definite estimates of this range are.

Method. Twelve standard slides showing unfamiliar foreign aircraft were used in the experiment selected from a large number available. All were identical in picture size and were projected so that the rectangular area thrown on the screen was 28 inches wide. All backgrounds were clear, and nothing was visible but the photographic image of an airplane. There was nothing but a figure within a frame. In four of the slides the airplane was large, occupying most of the frame (L-slides); in four others the airplane was medium in size (M-slides); and in the remaining four the airplane was small (S-slides). If the L-slides are considered the unit of size, the M-slides averaged five-eighths of this size and the S-slides one-half of it. Attitudes of the airplane were equivalent in the three types of picture. The twelve slides were presented for judgment in the order M, L, S, S, L, M, M, L, S, S, L, M. They were viewed at three distances from the screen, 10 feet, 20 feet, and 40 feet. If it is assumed that the "law of the visual angle" determines the apparent range of pictured aircraft, the range at these viewing distances will vary in the ratio 1:2:4. This was the assumption which had been tentatively made for the use of the new series of distant view slides.

Approximately 180 unclassified students awaiting entry into Preflight School served as subjects. They were divided into three groups of 60, one group judging the slides at each distance. They were taken 12 at a time and seated in a nearly dark classroom at the specified distance from the screen. They were instructed to make the best estimate of which they were capable of the probable

range or distance, in yards, of each of the aircraft shown. They were to guess if they were not sure. No other instruction was given. No practice slides were shown, and no standard of reference which might serve to stabilize the judgments was presented. They were untrained in the use of sighting devices, such as a ring-sight, for estimating range by means of the visual angle, and in any case no such devices were provided them or suggested. Estimates were recorded in writing. A total of 240 judgments were obtained for each of the three sizes of airplane at each of the three distances.

Results. The mean estimates of range in yards, with their standard deviation, are given in table 7.7. The variation in the

TABLE 7.7.—*Mean estimates of range in yards as a function of viewing distance and size in the picture*

Viewing distance	*L-slides*		*M-slides*		*S-slides*	
	M	*SD*	*M*	*SD*	*M*	*SD*
10 feet	206.2	213.4	313.8	224.2	413.8	289.3
20 feet	201.6	196.3	356.5	254.4	466.3	274.7
40 feet	234.5	214.3	407.2	265.7	556.0	276.9

estimates for each size of airplane at each distance is high, as is shown by the size of the standard deviations. The impression of distance produced by the pictures, insofar as the estimates are a measure of this impression, was evidently not a definite and clear one. The estimates tend to be guesses. The perception of distance in a picture without a differentiated background or reference points for depth might well be expected to be indeterminate. This interpretation was borne out by the experience of the experimenters. Nevertheless, the impression of distance was not wholly indeterminate. Reading the table horizontally, the mean estimates of range may be seen to rise in a consistent fashion. Evidently the impression of distance increases as the size of the plane, within the picture-frame, decreases. This fact also is in accordance with ordinary observation. As the measured size of the "large," "medium," and "small" airplanes decreased from unity to five-eighths to one-half the distance would be expected to rise in the proportion of 1:1.6:2 if there were a perfect inverse relationship. The mean estimates of range in table 7.7 actually rise (combining all three viewing distances) in the ratio of 1: 1.7: 2.2. There is, therefore, an approximately inverse relationship in aircraft recognition slides between the *size of the figure within a fixed picture frame* and the impression of distance.

Reading table 7.7 vertically, it is evident that apparent range did not increase in proportion as the viewing distance increased. If the range were determined by the visual angle principle and nothing else, it would be expected to increase in the ration 1 : 2 : 4. The obtained increases do not even approximate this. For L-slides the increase is approximately in the proportion 1: 1: 2. For M-

slides and S-slides the increase is in the proportion 1 : 1.1 : 1.3. These results indicate a tendency for the estimated range to remain constant regardless of the distance of the observer from the screen. The constancy is by no means perfect but it is at least clear that as the retinal size of a pictured airplane diminishes with viewing distance, its apparent range does *not* increase proportionally.

Conclusions. These results carry the suggestion that the use of pictures in instruction, when the pictures have to stand for real objects or situations, is subject to definite psychological conditions and limitations which are not always fully understood. The perception of distance in pictures is an especially complicated problem. The size of an airplane within a blank frame was shown to affect its apparent distance but the distance at which the picture is viewed did so only slightly. The impression of distance produced by such a picture, however, appeared to be subjectively indefinite and estimates of this distance varied widely. The implications of these results for training in aircraft recognition seemed to be as follows:

1. Students should not be expected to learn range-estimation of aircraft from pictorial material since the distance is more or less indefinite and is variable from one student to another. The 1,000-yard slides may look distant, but they will not necessarily look like aircraft at 1,000 yards.
2. Especially, students should not be instructed that the range represented on the screen will be greater as the distance of their seats from the screen increases.
3. The impression of an airplane at a distance can be produced by photographically reducing the size of the airplane within the frame of the picture.

A recommendation was added to these conclusions bearing on the question of whether it was desirable to produce special slides for training in aircraft recognition suitable for stereoscopic viewing in the classroom. Stereoscopic equipment could undoubtedly give an illusion of depth to the pictures shown on the screen so that the airplane would appear to be behind the plane of the screen, i. e., not at the same distance as the frame of the picture. It could also give an impression of three dimensionality to the airplane itself in a close-up view. The opinion of the experimenters was, however, that neither of these effects would be of any real value in training students for range estimation.

The Identifiability of Aircraft at Extended Ranges

Introduction. In view of the efforts to represent airplanes, in both still and motion pictures, at long range, it became a matter of some importance to know the smallest visual angle at which an airplane, in a given attitude, could be identified. The problem was

similar to that of measuring the so-called acuity of vision except that the stimulus object was an airplane.

This problem was obviously related to a broader one, a question of considerable practical importance to flexible gunners and to fighter pilots. What is the actual maximum range at which a given fighter or bomber can be positively identified in the air? How far away can one expect to be able to recognize a certain plane, and how is this related to the direction from which the plane is viewed? How does this maximum recognition range compare with the maximum firing range?

Actual measurement of this recognition range in the air, for all planes, and all attitudes of each plane, was obviously impracticable. A miniature experiment was possible, however, employing model planes and distances on the same scale as the models. It was at least possible to assume that the identifiability of a model seen against a background of sky at a considerable distance was approximately the same as the identifiability of a real plane seen against the same sky at a proportionately greater distance. The model and the real plane would intercept the same visual angle on the retina of the eye; both would be approximately silhouettes; and the details of both images would be equally just barely noticeable.

This assumption, however, was by no means certain. It will be examined in detail in a later section of this report, together with the factors which might modify it in one direction or the other. A safer assumption was that the *relative identifiability of different planes*, and of *different attitudes of the same plane*, when they are seen at long range, would be the same for models as for the real planes themselves. Therefore, using standard models of known scale, an experiment was undertaken to determine their identifiability at increasing ranges and in different attitudes. A discussion of the applicability of the results to the estimation of the absolute recognition range of real aircraft follows the experiment.

Method. In order to obtain data on the above questions, the identifiability of 6 model aircraft in each of 4 attitudes, seen against the sky, was determined at each of 6 distances varying from 12 to 98 yards. "Identifiability" was defined as the percentage of trained observers who could name the plane in the given attitude at the given distance. In order to insure that these percentages represented the actual recognizability of the models rather than the proficiency of the observers used in the experiment, planes were selected which should be known by 100 percent of the observers when the model was displayed at the nearest distance and in the most favorable view. This expectation was borne out by the results. A further check on the possible influence of proficiency on the results was made after the data had been tabulated.

Those observers who made any errors at all at the nearest distance were thrown out, and the results were recomputed. No appreciable differences were found in the values for the remaining distances.

The individuals used in the experiment were 307 aircrew trainees who had just completed a 30-hour course in Aircraft Recognition in the Preflight School (Pilot), Santa Ana Army Air Base. The aircraft models employed were the standard black plastic variety used in recognition classes. They were constructed on a scale of 1 to 72. Six planes were selected which were believed to be very familiar to preflight school graduates. Two heavy bombers (B-17, B-24) were used, two medium bombers (B-25, A-20) and two fighters (P-40, P-39). The models were mounted on the end of a wire holder about 10 inches long which was attached to a long pole. The pole was clamped upright to one edge of a large wooden platform. In this position the models were at a height of 13 feet from the ground, and were always seen against the sky at some distance above the horizon. The sky was a northern morning sky, neither brilliant nor dull, since the sun was approximately 90° away. The wire holder was fitted to a universal joint at the end of the pole, so that any desired attitude of the models could be obtained. The six planes were presented one at a time in each of the four standard attitudes: passing, plan, head-on, and "¼ front below" (i. e., halfway between head-on and passing, with the nose of the plane tilted upward 22°).

Each of four groups of about 75 cadets observed the models, one group for each of the four attitudes. These groups were divided into files of not more than 12 men. Each file began attempts to identify the models at a distance of 98 yards, and moved up to successively closer distances. At each distance the six models were displayed. This expedient helped to prevent earlier judgments from influencing later ones. It should be noted that the students were not aware that only six planes were being presented. The distances on the ground which served as observation points for each file of men were as follows: 98 yards, 73 yards, 49 yards, 37 yards, 24 yards, and 12 yards. On a scale of 1 to 72, these distances corresponded to 4, 3, 2, 1½, 1, and ½ miles respectively.

Each individual was provided with six cards on which to record his answers, one card for each of the six distances. He was given no knowledge of the stimuli except that U. S. Army aircraft only were to be presented. Since at least 18 responses were possible at each presentation, it can be assumed that a correct response involved a fairly positive identification of the plane represented. The instruction emphasized that wild guesses were not to be made. The planes were exposed for 15 seconds each in a prearranged order which was different at each distance. After responses had been recorded for all six stimuli at any one distance, each man in the

file handed in his card containing these answers. Comparing answers and talking were strictly forbidden during the course of the experiment, and enough proctors were provided for the enforcement of this condition.

Results. The percentages of correct judgments for each group on the six planes, and at the six observation distances employed, are given in table 7.8.

TABLE 7.8.—*Percentages of correct identifications of six aircraft in four attitudes at each of six distances*

Plane	*12 yards*	*24 yards*	*36 yards*	*49 yards*	*73 yards*	*98 yards*
GROUP I (N = 71)—QUARTER-FRONT BELOW						
B-24	100	100	99	100	99	98
B-17	97	100	96	99	94	83
A-20	100	96	97	82	23	32
B-25	100	99	96	99	44	17
P-39	87	62	54	51	41	17
P-40	87	56	55	27	11	8
GROUP II (N = 79)—HEAD-ON VIEW						
B-24	97	100	97	99	84	68
B-17	96	95	97	96	84	46
A-20	91	65	46	44	32	20
B-25	96	100	97	89	47	28
P-39	57	43	30	28	25	11
P-40	27	11	8	5	4	4
GROUP III (N = 71)—PLAN VIEW						
B-24	99	100	99	100	97	92
B-17	97	89	90	85	99	92
A-20	90	99	92	92	68	13
B-25	99	100	83	68	35	13
P-39	72	58	48	24	13	25
P-40	56	31	39	24	3	4
GROUP IV (N = 86)—PASSING VIEW						
B-24	99	98	94	91	72	42
B-17	100	98	98	99	93	83
A-20	100	90	73	58	30	30
B-25	100	91	67	41	13	13
P-39	99	93	83	66	16	10
P-40	100	98	81	52	19	10

The effect of distance is clearly indicated for the passing view of the P–40 (Group IV, bottom line). Though recognized by all 86 observers at 12 yards (scale equivalent of ½ mile), at 49 yards it is recognized by only 52 percent of the observers, and at 98 yards, the maximum distance, by only 10 percent. At this distance (scale equivalent of 4 miles) the *plan* view of the P–40 (Group III) is recognized by only 4 percent of the observers, whereas the plan view of the B–24, with a three times greater wing span, is recognized by 92 percent of the observers in that group. The *head-on* view of the P–40 is so readily confused with that of other airplanes (P–39 or P–51) that even at 12 yards it is recognized by only 27 percent of the observers. Many other similar comparisons can be made. In general the values for the percentage of correct responses resemble those found for most psychophysical measurements of perceptual functions.

In figure 7.2 these values have been plotted for each plane in each attitude. These graphs may be read to determine the maxi-

mum distances at which the models can be recognized with near-perfect accuracy. For example, at a distance of 24 yards, the B–24 is recognized in any of the four attitudes by over 97 percent of the observers. At 73 yards, however, though recognition of the plan and quarter-front-below views remains about as good, the nose view is recognized by only 84 percent, and the passing view

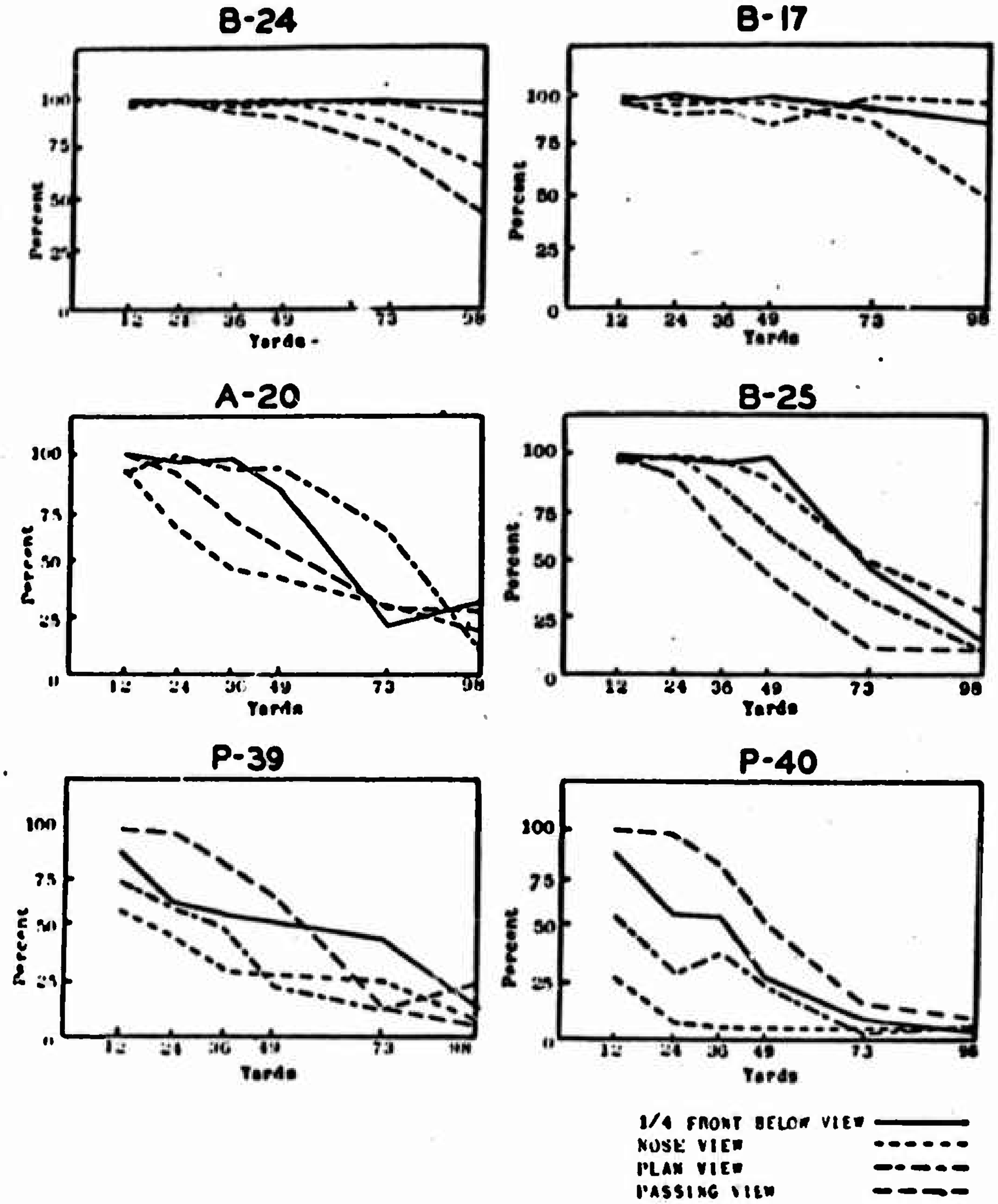


FIGURE 7.2.—Identifiability of Aircraft Models in Four Attitudes at Variable Distances

by only 72 percent. In the case of the P–40 and P–39, the passing view is the only one which is recognized by over 97 percent of the observers at 12 yards.

Other relationships in these data may be emphasized by comparing the distance at which the various attitudes of each plane are

recognized by an arbitrarily chosen percentage of the observers. The diagrams in figure 7.3 have been constructed by indicating the distance at which each attitude is recognized by 80 percent of the observers. The distance values were obtained by finding the intersection of each of the graphs in the previous figure with a horizontal line drawn from the 80 percent point, multiplying by 72, and converting yards to miles.

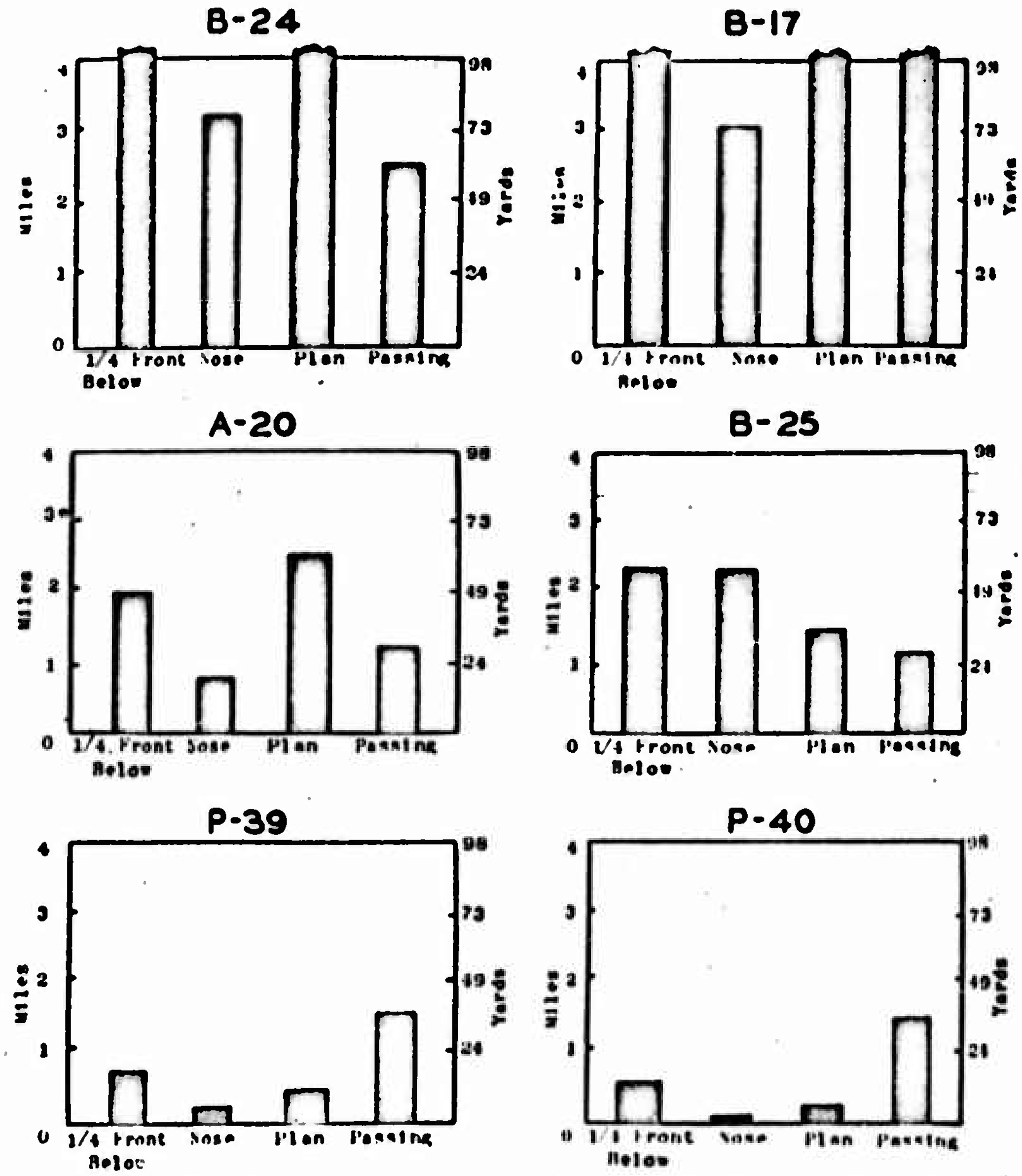


FIGURE 7.3.—Inferred Recognition Range, Defined as 80 Percent Recognition, for Six Aircraft in Four Attitudes

A number of conclusions may be drawn from the data depicted in these diagrams. Some of them are as follows:

a. The recognizability of a particular aircraft depends to a very striking extent upon the attitude in which it is seen.

b. For the small sample of planes studied, the most difficult of

703313 47—12

the four attitudes employed would seem to be the head-on position. That this is not always true is shown by results on the B-25, whose twin tails and slab-sided fuselage are easily seen from directly in front.

c. In general the attitude (of the four employed) which makes recognition at a distance easiest is the quarter-front-below view. If not the most recognizable, it is at least the second most recognizable view of each of the six planes studied. The B-24 is about as recognizable in plan view, the B-17 in passing view, and the B-25 head-on. The A-20 may be somewhat more recognizable in plan view. The P-39 and P-40 are decidedly more recognizable in passing view. It may be surmised that the reason for this lies in the necessity for distinguishing between the belly shapes of these two planes, a discrimination which can best be made from the side.

d. The identifiability of an airplane at a distance is seen to depend on its size. The two largest airplanes (heavy bombers) are the most identifiable at a given distance, the two smaller (light bombers) are next in order, and the two smallest (fighters) are least identifiable. The wingspans of these three types of aircraft are approximately in the ratio 3: 2: 1. This result is to be expected if identifiability, like acuity, is a function of the visual angle subtended by the object.

e. These data provide an explanation for many of the difficulties encountered by students in the recognition course employing slide views of aircraft. Illuminating comparisons are possible for instructors familiar with these difficulties.

Applicability of these Results to the Absolute Recognition Range of Real Aircraft. The validity of the distances given in figure 7.3 for combat recognition depends on whether the identifiability of a model seen at a distance is the same as that of a real plane at a proportionate distance. If the model is more identifiable, the ranges given in figure 7.3 are too great; if it should be less, the ranges are too small. No direct experimental evidence is available on this question. Indirect evidence, based on visual acuity research using various types of just-discriminable patterns would lead to the hypothesis that, under most circumstances, the identifiability of an object is about the same when it is small and near as when it is proportionately large and far. A 1-foot Snellen chart at a given distance is approximately as easy to read as a four-foot Snellen chart at four times the distance. But experiments of this sort have never been carried out at a distance (and size) which is 72 times the distance of the near object. It is not known whether the hypothesis would have to be modified at such long ranges. Moreover, the test objects used in these experiments (letters, broken rings, squares, gratings, etc.) are not the same as the shape of an airplane.

Acuity itself is not always an exact function of the visual angle of the test object. Under certain circumstances the Aubert-Forster phenomenon is reported which consists of a slight reduction in acuity for the large-far object. Under other circumstances, however, the phenomenon is reversed and acuity is slightly *greater* for the large-far object. But the distances employed in these experiments were not comparable to those with which we are concerned, and hence it would be hazardous to generalize from one situation to the other.

There is one factor, however, not present in the acuity experiments, which will operate unequivocally and in only one direction in the open air with real airplanes. This is the factor of atmospheric haze and the resulting diffraction of light waves. It will tend to reduce the identifiability of a real plane as compared with that of a model since, although the objects subtend the same visual angle, the rays from the real plane will have been diffracted by 72 times as great an amount of atmosphere. This will be true even on a clear day; and, of course, as atmospheric conditions of visibility worsen, the recognition range of real aircraft will decrease very markedly. Glare and the likelihood of a non-optimal brightness difference between plane and sky (the models were black whereas planes are silvered), are additional factors to consider. Taken together they make it very unlikely that the recognition range of real aircraft under ordinary conditions is greater than the recognition range as determined with models. Since conditions of visibility in combat are seldom ideal for recognition, the range figures given in figure 7.3 based on short distances and near-ideal visibility, will most probably be generally too high. Frequent conditions will operate to reduce these estimates; only rare or questionable factors will operate to increase them.

Conclusions. The question of what the recognition range of aircraft against the sky actually is has been shown to be a much more complicated question than discussions of it had assumed. In favorable attitudes of large planes it is very great; with unfavorable attitudes of small planes it is very small. The distance at one extreme may be at least 10 times the distance at the other extreme, if the data of figure 7.3 can be trusted. The importance of perceptible identifying features is again emphasized by these results. The suggestion is that these features become indistinguishable at varying distances. If the inferences of the last section are correct, however, it may be concluded that *under favorable circumstances* aircraft in general may be identified at distances greatly exceeding the firing range.

It is an interesting fact that the values obtained for identifiability of shape are related to distance in a similar way to those shown by other perceptual functions. There exist presumably

many types of visual acuity, of which this might be considered a special form.

SUMMARY

1. The use of rapid flash exposures did not prove to be of any practical value in learning to identify aircraft if the same pictures were shown with equal frequency for intervals of one second.

2. There was some evidence that verbal description of the visual shapes to be identified was a help rather than a hindrance to learning.

3. Training in reading digits or in estimating the number of spots presented in split-second exposures did not foster proficiency in identifying aircraft, nor did it improve the general efficiency of vision in terms of the criterion employed.

4. The ability to visualize the shapes of aircraft is related to the ability to identify them correctly. It may be concluded that the acquiring of differential responses is correlative with the acquiring of differential memory images. There was a tendency for these remembered shapes to be caricatures rather than copies of the objects.

5. Evidence was obtained that learning to identify these shapes was facilitated when differential reinforcement of the responses was provided.

6. The process by which the similar shapes of aircraft become distinctive probably involves, and would be aided by, a conceptual organization of the various shapes according to their significant features.

7. The size-distance relationship in viewing pictures or models of aircraft is complex. Plausible assumptions regarding it are not always correct.

CHAPTER EIGHT

Pictures As Substitutes for Visual Realities*

In the tests devised by the Film Unit and in most of the research on identification of aircraft, an underlying problem was repeatedly encountered, the problem of the appropriate use of *pictures*. The majority of these pictures were photographic—a category which includes both still pictures and motion pictures—but some were artificial in the sense that they consisted of nonsense shapes or highly schematic objects. These also could be either still or moving. Whether they were viewed on a projection screen or were seen in the form of photographic prints or drawings, they all had in common the characteristic of possessing a rectangular frame which filled only a small part of the observers' total visual field. They also were characterized by the fact that the screen or the paper on which the picture appeared was inevitably flat, and could be seen as such by the observers. These two characteristics may be taken as a preliminary definition of what is meant by a picture.

Pictures in general afford one method of setting up a "miniature reality." For teaching, training, general communication and entertainment they rival language in importance. They are easier to apprehend than language and presumably are perceived more directly. But the apprehension of pictures has its own rules which are different from those which apply to the understanding of language. And likewise the perceiving of pictures is governed by a different set of psychological conditions than is the perceiving of the situations represented by them. The visual situation represented in a picture, whether still or moving, is not only shown "in miniature," i. e., ordinarily reduced in size, but is subject to other differences and limitations. These differences become important when pictures are employed for exact purposes in psychological tests or for controlled types of training.

As described in the foregoing chapters of this report, a variety of scenes have been represented by pictures in the course of test construction and research. The following examples may be listed: the scene of distant airplanes flying through the sky at different

*This chapter was written by the editor.

velocities; a schematic instrument panel; the shapes of aircraft against the sky, both moving and motionless; the scene showing locomotion of the observer during a landing; the scene showing flight over the ground during a series of changes of direction. In the chapters to follow, even more complex scenes are required to be represented, such as the distances of objects in the third dimension, and the task of sighting a hand-held machine gun on an attacking enemy pursuit plane. A number of principles were found to operate in the presenting of these pictures to observers for testing or training purposes, which may be brought together briefly in the present chapter. They fall under two main headings, the equivalence of a picture viewed at different angles and at different distances, and the limiting differences between a picture and the scene represented by the picture. With respect to viewing conditions, there appear to be fewer limitations on the use of pictures than optical principles would lead one to suppose. With respect to pictures as substitutes for natural vision, however, there are more limitations than are often recognized.

THE EQUIVALENCE OF A PICTURE VIEWED AT DIFFERENT ANGLES AND AT DIFFERENT DISTANCES

The data presented in Chapter 4 demonstrated that within limits, the *performance* of an observer taking a motion picture test was independent of the angle at which he viewed the picture and the distance at which he viewed it. No direct evidence was available in these experiments with regard to the *appearance* of the picture. The most likely interpretation was that it had an equivalent appearance despite changes in the retinal image produced by different viewing angles and distances. The question of practical importance is this: does the *scene represented* in the picture become modified when the picture is viewed at an acute angle or at an "unnatural" distance?

Constancy of Representation at Different Angles of View. Ordinary experience in viewing photographs or pictures suggests that there is a considerable degree of latitude in the angle which the line of sight can make with the picture without distortion. It is known that the *shape* of an object seen in a straight-front position is preserved when the object is tilted or turned; this fact is given the name of "shape constancy." Presumably a picture, and the scene within the frame of the picture, are governed by an extension of the same fact of perception.

In order to study this principle, a simple experiment was performed and repeated on various occasions with different observers, employing a projector, a screen, and slides showing several types of scenes. The setup is diagrammed in figure 8.1. The normal arrangement of projector and screen is shown on the left. An

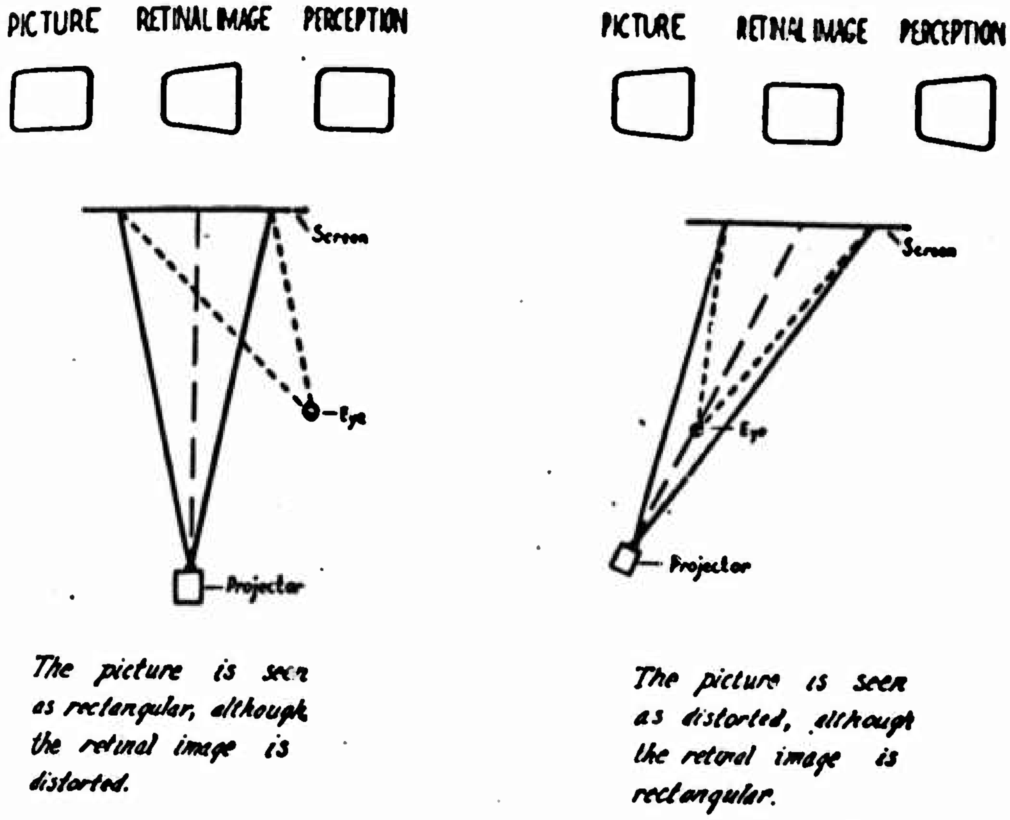


FIGURE 8.1.—The Effect of Oblique-View Compensation in Viewing Projected Pictures: Perceptual Shape-Constancy

observer viewing this screen picture at an angle reports that the *picture frame is rectangular,* and *the scene represented is not distorted,* although his retinal image is non-rectangular and foreshortened. The explanation of this fact must lie in the simultaneous perception of the screen as oblique. The theory suggested is that there exists an automatic compensation in visual perception for viewing at an oblique angle.

As a partial test of this theory the arrangement shown on the right was tried out, the projector throwing an oblique image. In a head-on view, the picture projected on the screen now appears foreshortened and distorted (the retinal image being distorted in the same way as before). But it also looks distorted when the eye is moved over to the same oblique point of view taken by the projector, the retinal image in this position being completely rectangular. When both projector and eye are on the same line of projection, optical distortion disappears. The critical variable in these different situations is apparently the perception of the screen as an oblique surface. The oblique-view compensation seems to be able to preserve the "constancy" not only of a normal picture but also of a distorted picture. This interpretation is consistent with the observation that when one looks at the arrangement shown on the right with half-closed eyes, or tries to see the picture

as an image dissociated from the screen, the distortion diminishes and its appearance comes close to being rectangular.

The implications of these results, although incomplete as a formal experiment, suggest that the viewing of pictures is governed not by the laws of optics taken by themselves but only as they are modified by principles of space perception. The principle of spatial constancy seems to be applicable.

Constancy of Representation at Different Distances. The scene represented in a snapshot or picture appears to the ordinary observer to be relatively unaffected by the distance at which it is held or viewed. Likewise a photograph may be enlarged or reduced in size without any obvious change in its capacity to represent a scene. There are probably limits to both of these generalizations, but the limits have apparently not been determined. People unquestionably have *preferences* in the viewing distances they select voluntarily for looking at photographs, musum pictures, and motion pictures. The Film Unit was unable to discover any empirical study of the basis of such preferences, or whether they are consistent. There is, however, an assumption which is accepted by photographers and is emphasized in the literature of photography that there exists only one proper distance at which a photographic picture should be viewed. This distance is *that at which the visual angle subtended by the picture at the eye of the observer is just equal to the visual angle subtended at the camera lens by the scene registered.*[1] In other words, the viewer must take such a distance that the proportion of the picture to his total field of view is the same as the proportion of the scene registered by the camera to the total field; the eye must be at a viewpoint equivalent to where the camera was. This requirement is said to be necessary if the distance and the relative positions and dimensions in the three-dimensional scene photographed are to be correctly represented and if they are to appear natural. It should be noted that, according to this rule, if a picture has been increased in size by photographic enlargement or by projection on a screen, the unique viewing distance must be increased in proportion to the degree of enlargement. This rule is not observed in practice, as indeed it could not be, in the viewing of motion pictures and in the showing of slides by projection to an audience. All members of the group cannot be seated at the "natural" viewing distance. The question is troublesome, and the problem arises as to whether the rule *ought to be* observed in the interests of correct representation.

The evidence accumulated by the Psychological Test Film Unit on this question has already been presented in Chapter 4 and in the experiment on the apparent range of represented aircraft in

[1] Henney, K. and Dudley, B. (ed.), *Handbook of Photography*, New York: McGraw-Hill, 1939, pp. 94-96.

Chapter 7. In the latter experiment there was a tendency for the apparent distance from the observer to the airplane pictured, indefinite as it was, to be the same whether the picture was viewed at 10 feet, 20 feet, or 40 feet. This result was in contradiction to the rule of the unique viewing distance. In these pictures, however, there were few cues to the perception of distance apart from the relationship of the object to the frame. Other pictures possessing perspective and representing three-dimensional scenes were subsequently observed under similar circumstances. The perspective, distance, and relative dimensions of these scenes did not appear to be distorted even though very considerable departures were made from the rule of the natural viewing distance.

The rule states that the observer of a picture must duplicate the visual angle of the camera that took the picture. Otherwise, it implies, his retinal image will not be a projected copy of the image which a spectator of the original scene would have had. But there is no proof that his retinal image needs to be such a copy. The visual surroundings of the picture-viewer consist of the room in which the picture is shown; the visual surroundings of the original spectator consist of the unrepresented parts of the total scene. These are quite different surroundings, and it is not legitimate to apply the optics of one situation to the optics of the other.

The evidence on the effect of viewing distance, together with other observations, suggested a perceptual rather than an optical theory of viewing pictures. In the case of photographs of three-dimensional scenes the *standpoint of the observer is itself represented in the scene.* The location of the observer in the space portrayed may not be exact, but it is never wholly indefinite. It depends, for one thing, on the amount of foreground visible (*cf.* chapter 9). The location of the observer in this space is something entirely distinct from his location in the space in which the picture is shown—the classroom, theater, or photographic salon. The location of the observer in the represented space, so long as he "loses himself" in it (or more accurately, *sees himself* in it) seems to be little affected by his location in the room-space which contains the picture. The picture itself is perceived in a substantially equivalent way whether its retinal image be relatively small or relatively large. This result may be explained on the basis of perceptual constancy. It is presumably for the above reason that the space represented in the picture is substantially equivalent whether its retinal image be relatively small or relatively large.

In all probability, too great a strain can be put upon this compensating ability of visual perception at extreme departures from the "natural" visual angle subtended by a picture. The appearance of photographs or motion pictures taken with telescopic lenses and the appearance of pictures taken with extremely wide-angle

lenses indicate that distortions of representation in the third dimension do occur. The limits at which these distortions begin to assert themselves are not known; the conclusion for present purposes is simply that the tolerance in allowable viewing distances for pictures is considerable.

THE SCOPE AND LIMITS OF PHOTOGRAPHIC REPRESENTATION

The attempt to represent complex spatial scenes for purposes of testing or instruction encounters both opportunities and difficulties. Photography is undoubtedly the most powerful method of accomplishing such representation, especially if one includes motion picture photography and the animation of drawings. Many striking and realistic effects can be achieved. But the differences between photographic representations and the process of seeing directly are nevertheless considerable. The visual and other mechanisms of man which yield his perception of a spatial world are, although subject to defects, superior to the mechanisms of representation by camera and picture at its best. In the course of the Film Unit's research, a tentative set of generalizations was gradually developed concerning the differences between the view yielded by the human visual mechanism and that yielded by photographic representation. In constructing tests for perceptual functions, it became evident that certain aspects of perception could, and others could not, be represented by pictures. Motion pictures added enormously to the possibilities but certain basic limitations were still present. A camera is capable of a good many kinds of "seeing," and a motion picture camera is capable of even more kinds. But no camera is capable of seeing as the eye sees. Inasmuch as some of the capacities of natural vision are not as obvious as universal possession of them might suggest, the differences and similarities are worth pointing out.

Angle of View

The field of view of the eyes is very much wider than that of the usual type of camera. Photographic representation has a narrow field of view, seldom exceeding 45° to 50° laterally. It therefore lacks the feature of peripheral vision, and consequently a picture is necessarily seen with surroundings which are extraneous to it. The only exception to this rule is panoramic projection of images, which may be passed over. What the observer of a screen of a photograph sees "out of the corner of his eye" is not the scene of the picture but the room in which it is shown. This rule becomes significant when it is desired to represent, for example, the view of a flier over terrain in which he must locate landmarks. The narrow field of the camera is a handicap which has to be taken into account.

The Effects of the Picture Margins

It follows from the above considerations that a picture must necessarily have margins. It cannot exist simply as a field of view; it is always an object within the field of view and therefore has a contour, which is conventionally a rectangular frame. The frame of a picture exercises a profound influence on the scene represented. Since the frame is always aligned with the vertical and horizontal directions of the room in which the picture is viewed, it becomes (or tends to become) the "frame of reference" with respect to which the orientation of the scene is judged. In the case of still pictures, if an object is not aligned with the frame, the object appears tilted, even though in actual fact it may have been the camera that was tilted. Even a scene showing terrain, horizon, or buildings tends to look like a tilted world in such circumstances, or like a hillside if that interpretation is possible, because of the strong tendency for the frame of the picture to determine the vertical and horizontal axes of the space represented. For this reason the camera in ordinary photographic work must be held in strict horizontal alignment and pointed horizontally forward. Under certain circumstances the camera may be successfully pointed up or down, as in aerial photographs, but only if there are cues in the scene to the unusual orientation of the viewer and his line of regard.

In the case of motion pictures *showing locomotion of the observer over the ground*, during flight for example, the frame of the picture may also, within limits, be shifted from the straight-front posture. It may be represented as applicable to the airplane and not to the ground. The experience of diving and turning can be shown with some success; the ground does not tilt up or rotate but instead the observer dives or turns. Banking is more difficult to represent; since the picture frame tends to be horizontal, the horizon itself tends to rotate and the observer then becomes disoriented. All efforts to *show the observer himself in space* by means of motion pictures must take account of the fact that the picture frame is an artificial frame of reference which comes between the observer and the spatial world depicted.

The Weak Sense of Orientation in Pictures

If one contrasts the facts just described with the kind of viewing of which natural human vision is capable, the differences are striking. The eyes, unlike the camera, need not be aimed straight ahead and held upright. One can lean over, look up, or lie on one's side without the slightest tendency for the visual world to tilt or swing. The explanation presumably lies in the fact that our retinal images, unlike motion picture images, are accompanied by and compensated by a postural sense of the orientation of the head

and the eyes. The perceiver of a photograph or screen picture *is not necessarily made aware by it of the orientation of the camera that took the picture.* In ordinary vision the perceiver is automatically and immediately adjusted to the orientation of his eyes by the existence of another sense.

This is not to say that the camera, and especially the motion picture camera, cannot represent orientation at the point of view of the observer and the direction in which he is looking, but only that there are definite limitations on this capacity. The only cues available for it are those which are present in the visual scene itself, and they are frequently not wholly adequate. The amateur who attempts to photograph a skyscraper by pointing his camera up at it gets a picture which is an adequate reproduction of his own retinal image but which looks unnatural. The explanation is not that the perspective is exaggerated, as sometimes is suggested, but more probably that the sense of looking up is missing from the picture. The spatial frame of reference for the scene is no clearly indicated, and the building may seem to lean away from the vertical or otherwise look "queer."

The Incapacity of Motion Pictures to "Look Around"

Early motion pictures were made with a motionless camera which registered a scene analogous to a theatrical stage or a still photograph. The motion was confined to the action shown and did not extend to the camera itself. "Pan" shots and "dolly" shots were a later development. Although modern motion picture technique employs a camera which moves with some freedom from one character to another and from one part of a scene to another, the shift in the "view" of the camera is usually slow and of no considerable degree. If it is desired to represent an object or event outside the field of view of the camera, a cut is normally employed rather than a moving camera.

This is not the state of affairs with natural vision. The eyes perceive a visual scene by a process of scanning it, i. e., they move from one fixation point to another, sometimes over a wide angle, by saccadic eye movements. A peculiar feature of these eye movements is that they are extremely rapid and that, for reasons only partially understood, the scene does not blur or even appear to move as the eyes sweep across it. The changed orientation of the eyes is automatically sensed. If a motion picture camera is turned rapidly from one line of regard to another, the picture blurs. Its capacity to shift direction is therefore limited, and the process of scanning or looking around has no real counterpart on the screen.

When the motion picture camera is used to portray subjective experiences, as it is on rare occasions in commercial films and as it was employed to good effect in a few wartime training films, the

limits on the kinds of "seeing" possible for it become a matter of importance. Conceivably they could be extended by novel techniques. The use of a "fast pan" instead of a cut under special circumstances to approximate the experience of scanning a scene might provide such an extension.

The Absence of a Focus of Attention in Pictures

The natural visual field has the characteristic of possessing a center at which vision is clearest. This center corresponds to the fovea of the eye—the part of the retina best equipped anatomically for exact perception. The periphery of the visual field becomes progressively less clear and vision fades or ceases at extreme angles. The margins of the visual field are not abrupt, as is the frame of a picture, and they are little noticed.

In contrast to this a picture lacks a central focus. The viewer may concentrate on any portion of it at will, and it is organized not by a gradient of clarity from center to periphery but by principles of what the artist calls composition. The camera cannot fixate on a scene in the strict meaning of the term; it cannot narrow the attention to a single object or a single portion of an object. The only means which the motion picture medium has of simulating the effect of concentrating attention is the conventional sequence of a long shot, followed by a medium shot, followed by a closeup. The part of a scene to be examined is, in effect, progressively enlarged until it fills most of the screen.

Absence of Binocular Parallax in Pictures

A picture, although it represents a three-dimensional space and may do so very adequately, is always at the same time seen as a two-dimensional object. The flat appearance of a still picture may be minimized and perhaps even destroyed by methods such as viewing a photographic print or transparency through an enlarging lens. The effect is to enhance the apparent depth of the picture to a marked degree. But a single photographic picture, unlike natural vision, presents the same stimulus to both eyes and consequently lacks the type of depth perception attributable to stereoscopic vision. The significance of this fact will be discussed in greater detail in chapter 9. Stereoscopic pictures add the cue of binocular parallax to other depth stimuli already represented in the individual single photographs, but they still retain some of the characteristics of a picture by virtue of having only a limited field of view and possessing a rectangular marginal frame.

The Point of View of the Camera and the Location of the Observer

Most pictures which show a terrain or which have perspective give some indication of the location of the observer in the space

represented, and some give a definite indication. The position of "here" can always be made out, i. e., the point at which the ovserver is standing. This point is not actually in the picture, but it is in the space represented by the picture. It is usually, but not necessarily, at approximately the point where the camera was located. The focal length of the lens employed, however, has an effect on this apparent standpoint, the result of a telescopic lens being to move it forward into the scene and the result of a wide angle lens being to move it backward from the scene.

Change of location of the observer, or locomotion, is represented by motion pictures with a considerable degree of success. The visual stimuli which make this possible are discussed in detail in chapter 9.

CHAPTER NINE

Perception and Judgment of Aerial Space and Distance As Potential Factors in Pilot Selection and Training*

EVIDENCE THAT SPACE PERCEPTION IS IMPORTANT IN THE SELECTION AND TRAINING OF PILOTS

Aerial space may be defined as the visual surroundings extending away from the observer and bounded in any direction by the horizon, the surface of the earth and the sky. It may be distinguished from local space primarily by its voluminousness and long range of distances. Local space is the kind to which we are accustomed; it is inclosed by walls and restricted in range by them. Even out of doors in a civilized environment the spatial scene is cut up and confined to localized areas by buildings and other objects which obliterate the horizon. It is aerial space which constitutes the environment of the flier.

Persons who are adapted to going about and making the ordinary judgments of distance in the city are usually misled by the extent of distances in the desert, mountains, on water, or from a plane. Generally, aerial distances are poorly estimated by such persons because they are unfamiliar with the visual cues present in the situation for space perception. The spatial adjustments which are adequate in local or room-sized space are not adequate for flying a plane. It may be assumed that the pilot must possess or acquire the ability to perceive aerial space accurately.

Although indirect evidence and tradition in aviation medicine provide a basis for believing that aircrew personnel, particularly the pilot, must have accurate aerial space perception, semi-empirical evidence is also available. There are three sources of such evidence in the research carried out by aviation psychologists in the AAF: (1) Job analyses, which have been carried out to determine the essential abilities needed by the aviation student in order to learn successfully to pilot a plane in a short period of time; (2) Statistical or factorial treatment of instructors' grades

*This chapter was written by the editor in collaboration with N. M. Glaser.

on various flight maneuvers, which has yielded the primary factors which determine these grades; (3) Study of the causes for accidents during training, which has revealed areas of low aptitude.

A study of the reasons for the elimination of over 1,000 students in primary flight training was one of the first efforts to determine the abilities needed by the pilot. The Faculty-Board proceedings for these eliminees were reviewed and from them a list of the 20 traits most frequently mentioned was compiled. In 30 percent of the failures reported by the board at the primary-school level; faulty estimation of speed and distance was mentioned as a contributing factor. A rating scale of the 20 traits was constructed and another sample of 1,305 students eliminated from elementary pilot training were rated by their instructors.[1] The inability to estimate distances and velocities in tridimensional space was checked as a cause for elimination in 31 percent of the eliminated students.

The maneuver of landing a plane is considered an outstanding example in which accurate perceptual judgment is needed by the pilot. One analysis of the problem of landing[2] concluded that the principal difficulties encountered by primary student pilots were stalling-out correctly, placing the gliding turn correctly, maintaining a straight approach leg, and breaking the glide at the correct height. More specifically, learning to land a plane depended on the ability to learn and use visual cues for height, distance, direction, and velocity of motion in space.

The results of factor analyses provide quantitative evidence that spatial skills are present in operating an airplane. An analysis[1] of the intercorrelations between daily grades on various maneuvers yielded three rotated factors named: (1) Perceptual Judgment; (2) Headwork; (3) Motor Technique. Perceptual judgment had the highest loadings for spins, landings, traffic, and forced landings. It is defined as "the ability to make rapid and accurate judgments of distance, speed, and altitude."

Probably the strongest incentive for the construction of tests of space perception came from the reports of accidents in pilot training. One of these reports by the Field Studies Unit, Office of the Surgeon, in the Headquarters of the Training Command, tabulated the results of 180 accidents that occurred during elementary training and 97 accidents that occurred during basic and advanced training.[3] It was found that 20 percent of all accidents

[1]These studies are given in comprehensive report No. 8: Miller, N. E. (ed.), *Psychological Research on Pilot Training*, AAF Aviation Psychology Research Reports, No. 8. Washington: Government Printing Office, 1947.

[2]Data of this type may also be found in comprehensive report No. 8. The study was performed by Psychological Research Unit No. 3.

[3]*Ibid.*

in primary schools, and 23.8 percent of those in basic schools were connected with landing maneuvers. The investigation concludes: "The primary factor involved in this type of accident appears to be a lack of ability in the perceptual field, especially the ability to estimate rates of speed, and distances, and spatial orientation."

If the perception of aerial space and distance are of such importance for the selection and training of fliers, it is obvious that psychological tests and training methods need to be devised with which to select and train them. This cannot be done effectively until the nature of space and distance perception is understood. At the present time, most psychologists would probably agree that it is not adequately understood. It is the purpose of this chapter to provide a theory for a clearer understanding of space perception and to describe the experiments and tests involved in the theory. A systematic analysis of aerial space and distance should, if correct, have many practical applications to the problems of pilot selection and training.

THE TRADITIONAL PSYCHOLOGICAL PROBLEM OF DEPTH PERCEPTION AND THE EMPHASIS ON OCULAR CUES

The Assumption of the Binocular Basis of Depth Perception

If it can be taken as proved that the pilot has to be able to judge tridimensional space in order to fly successfully, what is the sensory basis for the perception of such space? This question is, of course, the ancient problem of how we see a world which appears to extend away from us rather than a flat world, analogous to a picture, corresponding to the image formed on the retina of the eye. The accepted answer to this question—the answer given in the literature of aviation medicine and also by most of the textbooks in psychology and physiological optics—is that depth perception has its basis primarily in the existence of two eyes. The fact of binocular parallax, or stereoscopic vision, is commonly referred to as the main explanation of depth perception. It is usually stated that the binocular cue is supplemented by "monocular" cues for the perception of distance, but these are usually thought of as secondary. It is supposed that these latter signs or indicators of depth are not innate but are learned in the course of experience and therefore have little to do with the pilot's intrinsic or essential ability to see depth. These monocular cues are usually listed as including such factors as linear perspective, transposition of objects, shadows and shading, aerial perspective, and occasionally a few others. They will be discussed in the next section. The question which arises here is whether the accepted emphasis on binocular vision is correct insofar as it concerns flying.

A good deal of evidence can be adduced to show that visual cues which are *not* dependent on the spatial separation of the two eyes

703313 -47- 13

are of much greater significance for the kind of distance perception which fliers need than has been realized in the past. The evidence will be listed in the following paragraphs.

The Perception of Distance by Persons with Only One Eye. There has been a sufficient number of monocular pilots who flew successfully to suggest that binocular vision is at least not absolutely essential for adequate flying performance. The most famous of these was Wiley Post, who was admittedly an excellent flier. If space can be judged successfully with the use of only one eye, then the monocular cues of the normal pilot with two eyes must also be capable of producing space perception. Probably the normal pilot has even better capacity for such perception because of the fact that each eye supplements the monocular vision of the other eye, quite apart from the binocular disparity of the two images, and because two eyes yield a wider field of vision than one eye alone. Training or experience may or may not be necessary for monocular space perception; the point is simply that the capacity is present.

The Perception of Depth in Photographs and Pictures. It is a familiar fact that depth perception can be produced artificially in the stereoscope, i. e. by presenting separately to each eye the picture which it alone would see in the corresponding real three-dimensional scene and superposing the two different pictures by prismatic lenses. The vivid perception of depth which results is taken to be a proof of the effectiveness of binocular or stereoscopic vision. What is less familiar is the fact that a striking depth effect can be seen if two *identical* photographs are substituted in the stereoscope for the two pictures taken from slightly different points of view. The depth effect in this case is frequently comparable to that obtained with genuine stereoscopic viewing. Similarly, if a single photograph of a three-dimensional scene is viewed in such a way as not to emphasize the flatness and the frame of the picture, the observer frequently gets as much effect as if he were looking through a stereoscope. As Schlosberg[4] and others have shown, the explanation is apparently that one sees depth in these single pictures because they are viewed through a lens which minimizes the surface quality of the picture and which hides its frame. This is the method by which stereoscopic photographs are viewed. The conclusion must be that a considerable part of the depth effect obtained with the stereoscope itself is not a genuine binocular effect at all but instead is dependent on the monocular stimuli for depth present in the single photographs but ordinarily inhibited by the circumstances under which they are viewed. These cues lose much of their effectiveness under the customary conditions for looking at photographs because they are contradicted by

[4] Schlosberg, H. Stereoscopic depth from single pictures. *Amer. J. Psychol.*, 1941 54, 601-608.

the cues which make the picture a flat rectangular surface. If the conclusion is valid for stereoscopic photographs it must also be valid for ordinary binocular seeing, i. e. it is implied that a considerable part of the depth effect in ordinary vision is not binocular but monocular in origin.

The Diminishing of the Binocular Cues with Distance. It is a possibility that aerial distance perception at long range is mediated somewhat differently from distance perception at short range and that while binocular cues are important in the latter situation, they are less important in the former. It is likely that the depth effect produced by binocular parallax becomes ineffective beyond a certain distance from the observer. The eyes are about two and one-half inches apart. For objects in near space, this is enough to produce parallax; or otherwise stated, there will be a disparity between the images in the right and the left eye, which serves as one kind of stimulus for the perception of depth. But for objects in far space, the retinal disparity in the two eyes presumably becomes so minute as no longer to be an adequate stimulus for seeing depth and, for all practical purposes, the two eyes have identical images. At just what distance from the observer this occurs does not seem to be agreed upon; the range of stereoscopic vision is sometimes given as under a hundred feet and by others is estimated at a distance of as much as a thousand yards. All such estimates seem to be based on calculations rather than on empirical measurement of the effect of disparate retinal vision in real space. They assume that the just-noticeable retinal angle of disparity as determined with a stereoscopic apparatus is the determining factor for the maximum distance at which one can still see binocular depth in the open air. The actual range of stereoscopic vision, therefore, is not known. It is fairly certain, however, that the other binocular cue of the degree of convergence (with correlative accommodation) of the eyes has a fairly short range. At longer ranges, both convergence and accommodation disappear. They are, of course, essential for normal vision but as criteria of distance they are limited to what has been called room-sized space and are ineffective for the perception of aerial space. The only conclusion that can safely be drawn is that since the effectiveness of the binocular cues decreases with distance, the monocular cues are probably increasingly significant at large distances, and may even be the only cues available at such distances. Presumably it is this long-range distance perception which is important to fliers.

The evidence above all points to the conclusion that the visual stimuli for depth *not* dependent on the spatial separation of the two eyes—the so-called monocular cues—need to be taken into account in selecting and training fliers for effective space perception.

The Monocular Cues for Depth Perception

The list of accepted cues for the perception of depth has very largely remained unchanged since the discovery of the stereoscope. The non-binocular cues are sometimes called signs or indicators or criteria of depth to imply that they have not the same status of elementary sensations as has the fact of binocular retinal disparity. They are conventionally thought of as having to be interpreted rather than being sensed and they are assumed to be learned rather than innate. The list usually includes some or all of the following factors: *linear perspective* (such as converging railroad tracks), the *apparent size of objects of known size* (which decreases with distance from the observer), the *changes due to atmospheric conditions such as haze* (aerial perspective and blurring of outlines), *monocular parallax* (change of appearance with change of the observer's position), *interposition* (the superposing of near objects on far objects), *shadow patterns* (the light-and-shade relations yielding relief) and sometimes the *angular location of the object on the ground* (position of the object on the retinal dimension beginning with the observer and ending with the skyline). *Accommodation* is also sometimes given as a monocular cue for near depth. It is evident that all of these cues are not on the same explanatory level. Some of them will explain not how the distance of an object is visible but only how one object can be seen at a greater distance than another. For example, interposition and shadow patterns give the *relative* location of objects but do not produce the impression of a space which is continuous in the third dimension. Although all these cues have been described by many observers, they have in general not been experimentally isolated or systematically varied in relation to the perception of distance. They are described somewhat differently by different writers and have not been brought together into a consistent theory explaining how they can function. Nevertheless, if they are as significant for the perception of distance by fliers as seems likely, it is important that such a theory be formulated. If they are to be used as a basis for tests of the ability to judge distance or if they are to be described with sufficient exactness so that they can be used in training, they must be redefined. An attempt to define them and to formulate a theory will be made. Before doing so, however, it would be well to look into the question of the kind of space which they are required to explain.

The Kind of Distance Perception Required for Flying

When one describes the cues for the perception of distance in the terms above, the perception referred to is the distance of a particular object rather than the impression of continuous distance. Conceiving the problem in the traditional way, distance

perception in general consists of the ability to judge the distances of a number of specific objects. This, however, is not the space in which the pilot flies. What he perceives is a continuous space. It is almost never a single distance which he needs to judge, but a dimension of distance. There is invariably beneath him a continuous terrain, and what he discriminates is the location of all points on this terrain rather than specific distances to given points. Objects on or above this terrain may be momentarily of great importance, it is true, but they are judged in terms of a continuum of distance or, in other words, a background of three-dimensional space.

Traditionally conceived tests for the perception of distance have concentrated on the problem of how well an observer can judge the relative distance of two objects, or how accurately he can equate the distance of two objects. But the judgments a flier makes are in terms of appropriate changes of speed and direction of flight in relation to the distance of the ground. Such distance judgments always involve the "here" position of the observer at one end of the distance to be judged. It might be suggested that the practical value of depth or distance perception is that it makes possible locomotion through a continuous space which includes obstacles, and that both the obstacles and the locomotion itself involve the absolute distance from here to there.

Tests for depth perception, therefore, should aim to set up a kind of judgment similar to this. And the theory behind it should be a theory of a continuous space with an underlying terrain in which the observer is himself located and in which he can move.

THE STIMULUS VARIABLES FOR THE PERCEPTION OF DISTANCE AND CONTINUOUS SPACE IN THE OPEN AIR

The problem of three-dimensional vision, or distance perception, is basically a problem of the perception of a *continuous surface* which is seen to extend away from the observer. All spaces in which we can live include at least one surface, the ground or terrain. If there were no surface, there would be no visual world, strictly speaking. Whether we stand on it or fly over it, the ground is the basis of visual space perception both literally and figuratively. It is obvious enough that we could not stand or walk without the ground, but it is equally true that a pilot cannot fly purposefully without the ground and its horizon to guide and orient him. If by reason of fog or darkness the ground is invisible, an instrument must be provided to give him a substitute for it, an artificial horizon. The terrain, of course, is not all there is to the visual world. Objects stand out against the ground and they are usually what demand our attention. But an array of objects by themselves does not make up visual space; it is constituted instead

by the ground or surface against which these shapes and figures appear. The visual world consists of object-surfaces on a background of an extended ground surface. This is what is implied by the "figure-ground" distinction in perceptual psychology. If we ask how the distances of these objects are seen and discriminated, it would be a mistake to disregard the surface of the background which connects and lies behind them. This mistake has regularly been made in most theories of depth perception. We need to explain not the "cues" or "indicators" to the distance of specific objects but instead the dimension or sensory continuum of distance, *as such,* which, once visible, determines how distant all the objects within it are.

This view of the problem is in contrast to the classical formulation which asks how the retina of the eye can see a third dimension in the sense of a theoretical line extending outward from the eye. Points on this line at different distances must all be identical so far as the retina is concerned. Nevertheless we do see depth. How can this be? The solution to this dilemma is to recognize that visible distance does not consist of a line extending outward from the eye. The question to ask is not how do we see such a line but how do we see the substratum—the surface which extends away from us in the third dimension? The image of this surface is obviously *spread out* across the retina.

Figure 9.1 illustrates the two formulations of the problem. The points A, B, C, and D cannot be discriminated by the retina. Distance along this line is a fact of geometry but not one of optics or of visual perception. But the points W, X, Y and Z at corresponding distances can be discriminated by the retina. They represent the retinal image which corresponds to an extended substratum. It may be noted that the retinal points become progressively closer together as the distance increases.

If this view is correct, it is necessary to see a continuous surface in order to have an accurate sense of continuous distance. The sky may be a background but is not a surface. Distance appears to end at the skyline and the sky itself does not have a determinate distance. Single aircraft or clouds in the sky are of course objects having a surface, but since there is no background surface behind them, their distances ought in theory to be difficult to estimate, and in actual fact they are.

The stimulus variables which make possible the perception of such a continuous surface must necessarily consist of continuous differential stimulation on the retina. The retinal image of the surface must differ significantly at different points corresponding to those which are farther or nearer. There must, in other words, be retinal *gradients* of stimulation. The present use of the term "gradient" may be explained by the following illustration. It is

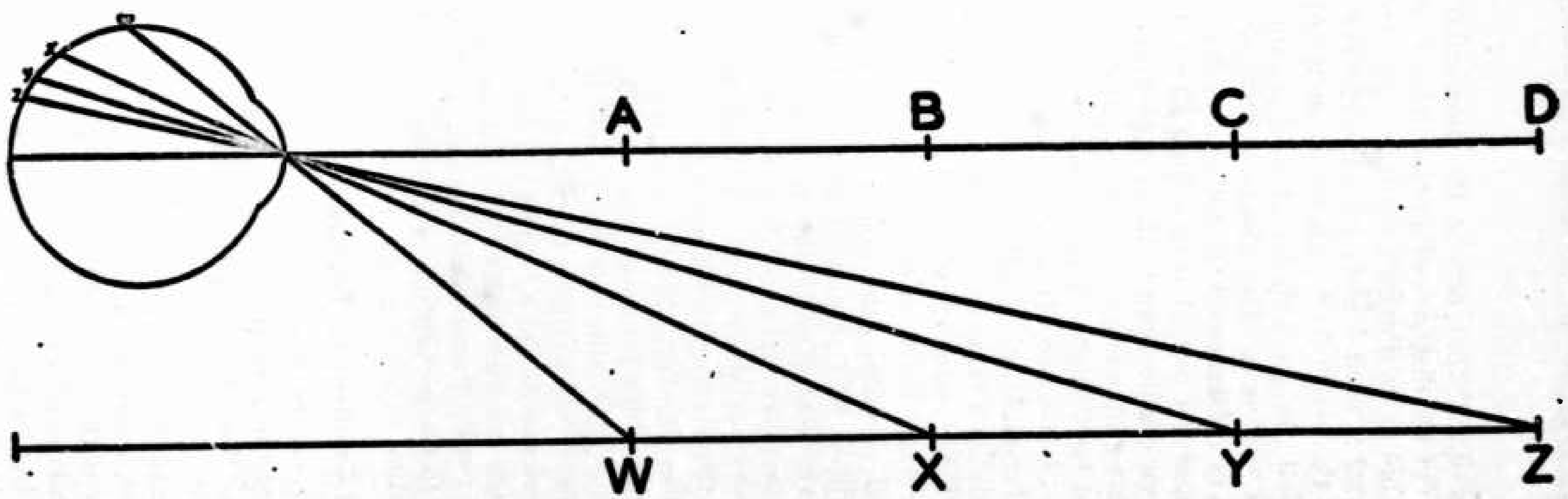


FIGURE 9.1.—Two Formulations of the Problem of Distance Perception.

sometimes stated that one of the monocular indicators for the perception of the distance of a point in space is its retinal location on the up-and-down dimension which begins with the lower margin of the visual field and ends with the horizon. Usually the lower margin of the visual field includes an image of the observer's feet and body—it always includes at least a faint marginal image of his cheeks and nose. The observer himself and the skyline are two points of reference on the retina and the distance of the object from the observer may thus be estimated on the basis of its visible up-and-down relationship to these two points of reference. Let us consider this statement. It is very doubtful if this retinal dimension should be thought of as a sensory variable *as such* for the perception of distance. It would be a stimulus only if there were differential stimulation yielding an extended *surface* in perception. The up-and-down location of a retinal point has a distance value only when it is located in relation to a gradient of retinal stimulation. The retinal limits of the skyline on the one extreme and the "bottom" of the field (the body) at the other are limits within which gradients of stimulation may lie, and, as we have already implied, a gradient of stimulation must exist if a continuous distance is to be perceived extending into the third dimension.

The sensory variables which underlie the perception of distance as defined above can now be described. The list will be found to differ considerably from the familiar list of cues for depth perception. The variables proposed are intended to be genuine dimensions of the stimuli affecting the retina, like the stimuli for color and brightness, and to differ from them chiefly in that the dimension is spread across the retina in the form of a gradient and that it is of a more complex order. To what extent they are learned or innate need not be discussed at this stage. They are all systematically related to the perception of a continuum of distance embodied in a substratum extending out to the horizon.

The Retinal Gradient of Texture

The difference between the perception of a surface, such as a flat wall, and the perception of an area without surface, such as the sky, has been investigated in the psychological laboratory. According to Metzger[3] and also Koffka[4] the difference lies in the fact that the surface corresponds to a retinal image having minute irregularities, spots, or differences in stimulation from point to point, whereas the area without surface corresponds to a retinal image which is in effect completely homogeneous. The area is differentiated in the former situation and undifferentiated in the latter. The term which Metzger and Koffka use for this sensory

[3] Metzger, W. Optische Untersuchungen am Ganzfeld. II. Zur Phänomenologie des homogenen Ganzfelds. *Psych. Forsch.*, 1930, *13*, 6-29.

[4] Koffka, K., *Principles of Gestalt Psychology*, New York: Harcourt-Brace, 1935, ch. 4.

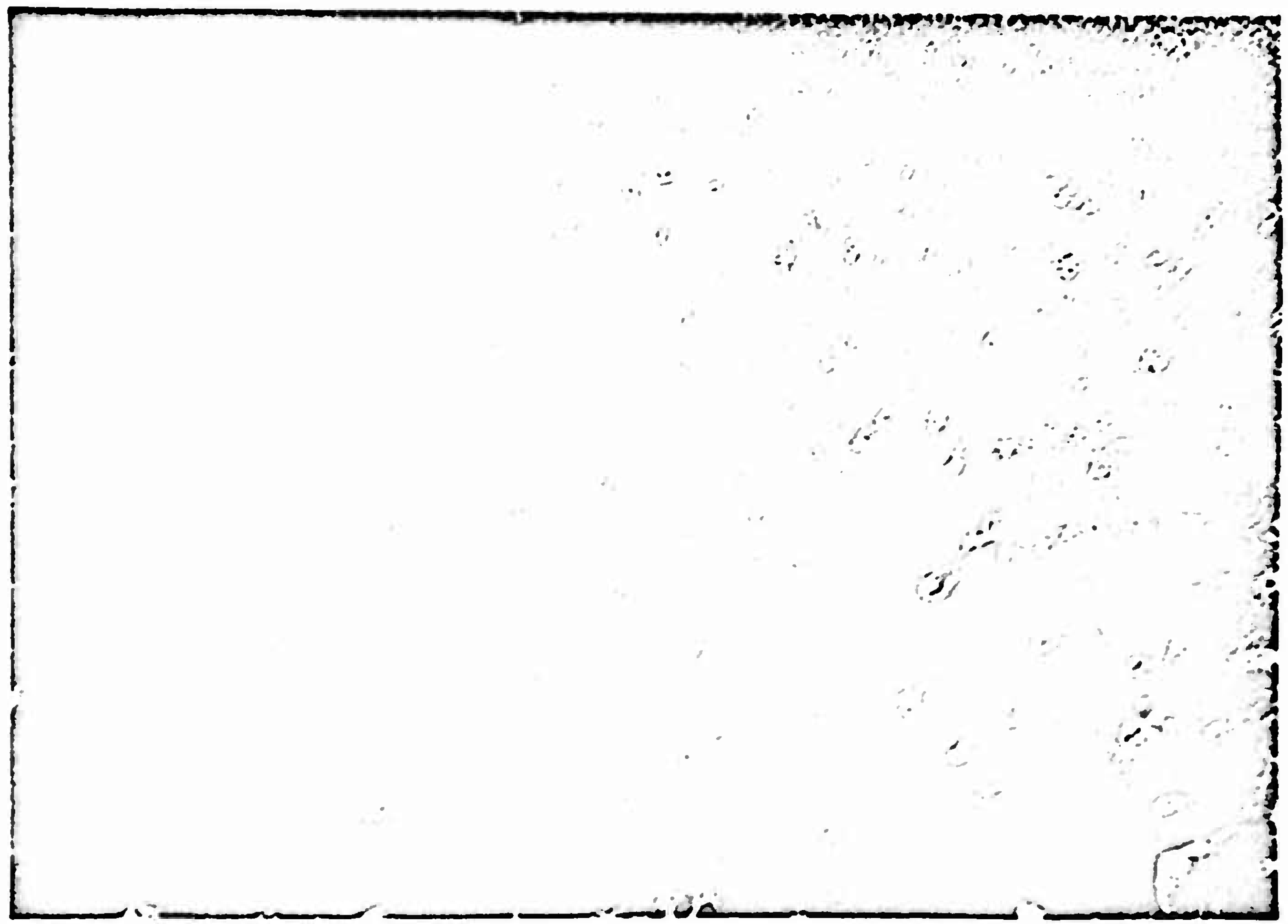


FIGURE 9.2.—"Distance as Produced by a Natural Gradient of Texture."

quality is *microstructure.* When an area of the visual field has microstructure, a surface is visible at a determinate distance; when the area has no microstructure, nothing is seen but "film-color" and no determinate distance is visible.

It is possible to go a step farther and to point out that the retinal image may vary between extremely coarse and extremely fine differentiation. In order to include the extremes of this stimulus variable, it will here be called not microstructure but "texture." As a first approximation to a definition, it may be suggested that retinal texture is the size of the "spots" and of the gaps between them in a differentiated visual image.

Any surface, such as the ground, obviously possesses texture. If it extends away from the observer, the retinal texture becomes finer as the distance of the corresponding points of the surface becomes greater. Figure 9.1, already discussed, indicates the way in which the retinal image becomes more "dense" as one passes from point W to point Z. There will exist a continuous gradient of texture from coarse to fine with increasing distance of the surface. A retinal gradient of this sort is in fact an adequate stimulus for the perception of continuous distance whether or not it is produced by an actual surface extending into the third dimension. The effectiveness of this stimulus-variable may be illustrated by three examples. In figure 9.2, there is a gradient of texture from coarse to fine running from the bottom to the top of the picture and, correspondingly, a continuous increase in the visible distance of the surface. In Figure 9.3 the same effect may be seen but with a texture of different character, i.e., a texture having elements of different shape and different mean size. The *gradients* in both pictures are, however, similar. It is an incidental fact that these texture-gradients were produced by photographing a ploughed field in the first illustration and a stubble field in the second; it is nevertheless true that the only effective stimulus for distance perception in the pictures is the variable of texture. Figure 9.4 may appear to be an even more convincing demonstration of the stimulus variable, since the gradient of texture was here constructed artificially. The line segments in this illustration were drawn increasingly smaller from the bottom to the top of the picture and so likewise were the vertical and horizontal spaces between them. The impression of a level terrain extending away from the observer is compelling.

It may be noted that the stimulus-correlate of distance in these illustrations is not the gross retinal size of the texture-elements but their *relative* size within the gradient. For example, the size of the line-segments in Figure 9.4, i.e. the elements of the texture, could have been twice as large at the bottom of the picture and would then have been twice as large all the way up the picture to

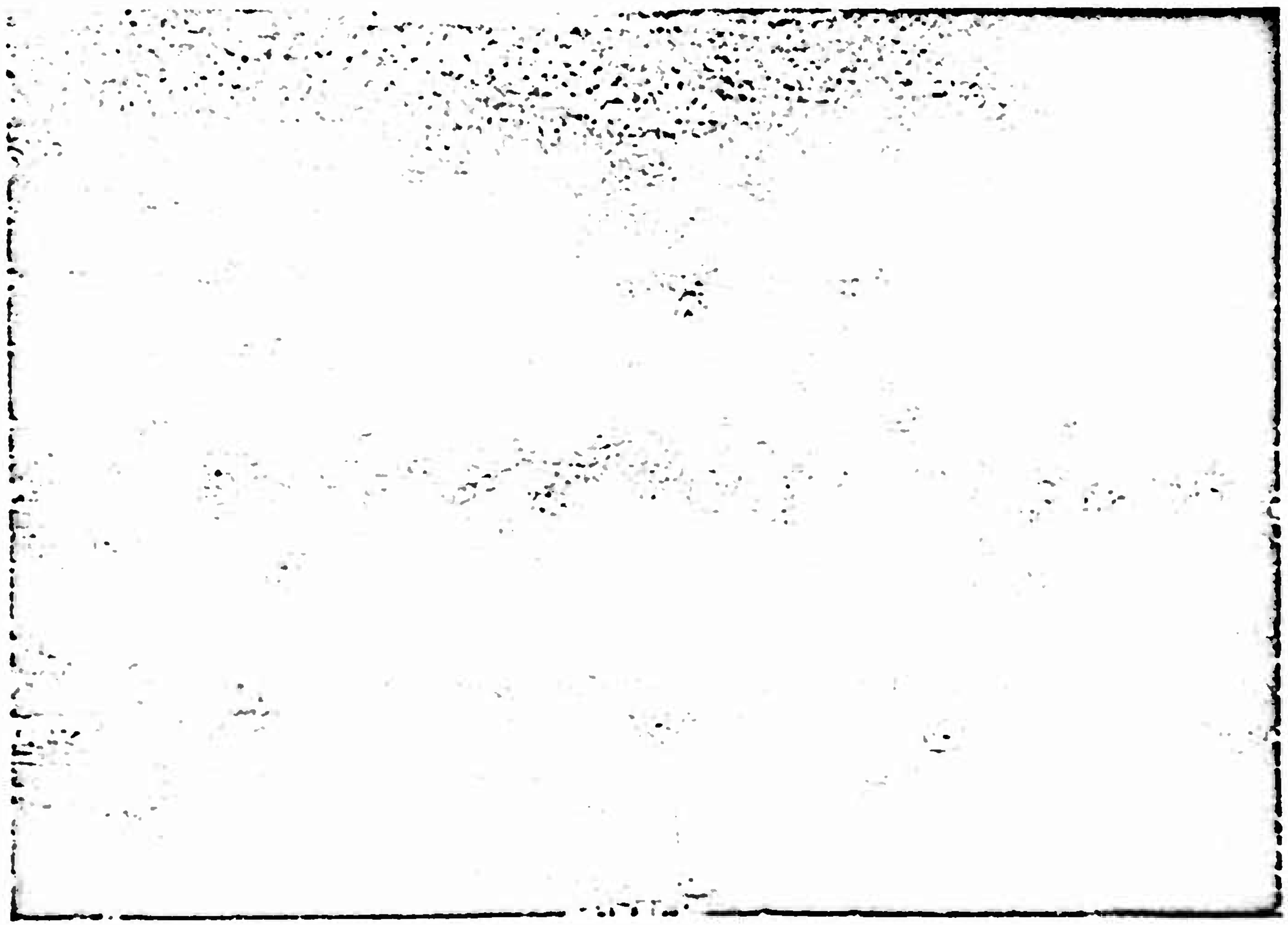


Figure 9.3.—"Distance as Produced by a Different Natural Gradient of Texture."

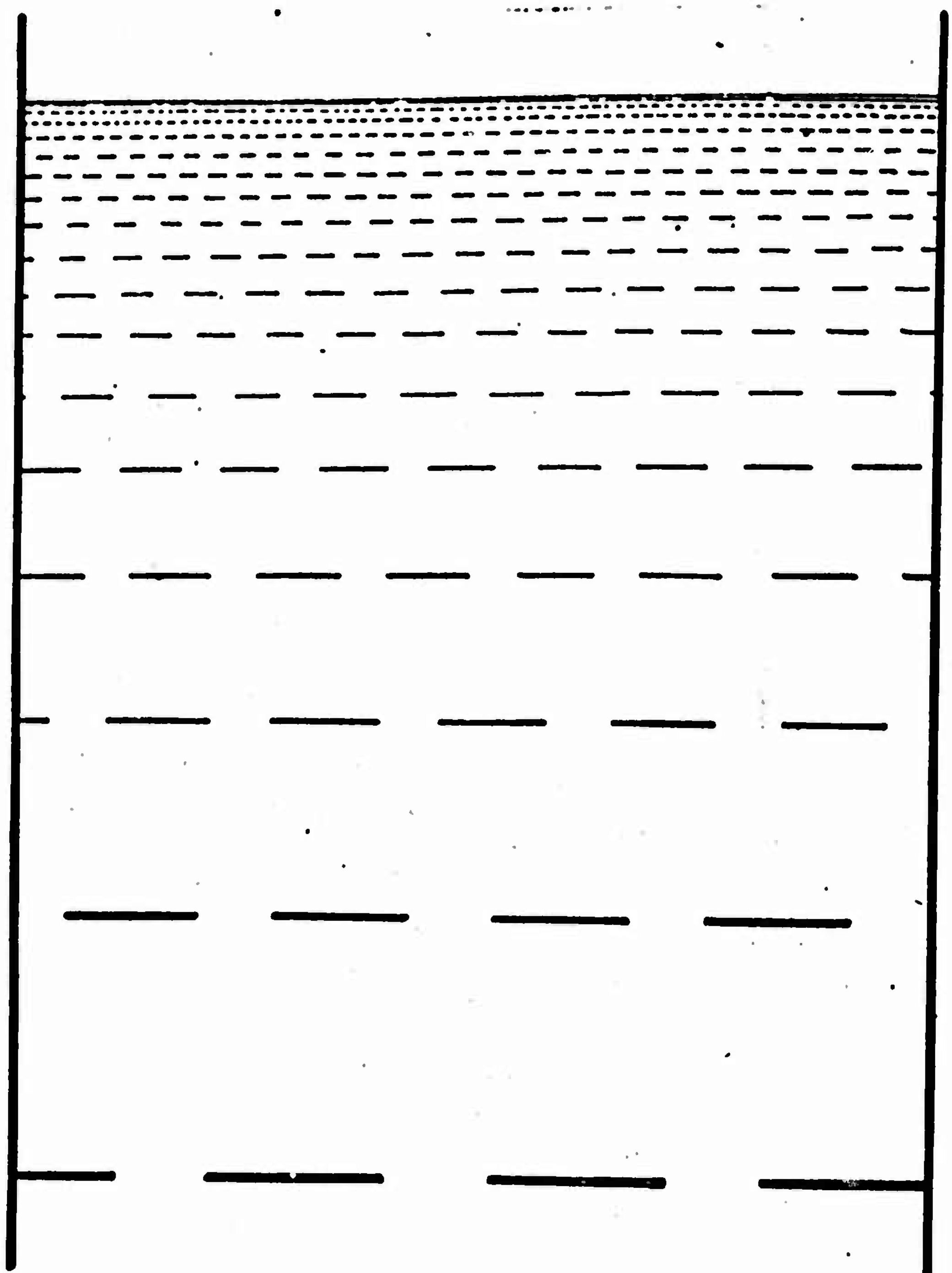


FIGURE 9.4.—Distance as Produced by an Artificially-Constructed Gradient of Texture.

the horizon; the resulting impression of distance, would, however, have remained the same. This implies that in perceiving distance over real terrain, it is a matter of indifference to the observer whether the over-all texture of the terrain is made up of large or small elements — whether for example it is produced by sand, grass, brush, or trees.

It should also be noted that the line elements of Figure 9.4 were so drawn as not to fall one behind the other in straight lines con-

verging to the horizon. This would have introduced the factor of linear perspective, which ought to be considered separately. The stimulus variable in that illustration was intended to be one of "pure" texture. The texture gradient is, however, a kind of perspective in the broad sense of that term and it is related to linear perspective inasmuch as in the case of both variables retinal size decreases with distance and vanishes at the horizon. All the retinal gradients to be described as stimulus variables for distance perception are analogous to perspective at least in respect to being extended on the surface of the retina. The variable just described, therefore, might well be given the name of *texture-perspective.*

The Retinal Gradient of Size-of-Similar-Objects

In almost every kind of terrain which the flier is likely to meet, and in most of the spaces of everyday life, objects are present in addition to the substratum itself. Commonly there are classes or types of similar objects scattered about or lined up in the environment. Houses, fence posts, telegraph poles, fields, and even hills tend to be of similar physical size and shape, as do chairs, tables, and people. If there are more than a few of these similar objects in the visual field, there can exist a gradient of decreasing retinal size corresponding to their distance from the observer. The principle involved is the familiar one of size perspective. If more than one *homogeneous type* of object is present, there will be more than one gradient, and it may be assumed that different gradients may exist at the same time such as, for example, one for trees and another for houses. Gradients of size and gradients of texture are obviously analogous and the one merges into the other when the objects in the visual field become sufficiently numerous.

If the objects on a terrain are lined up in rows, or if extended objects like roads and fields having linear contours are present, the size perspective becomes linear perspective. This stimulus for distance is more familiar than the others, but it is merely a special case of the principle that retinal size decreases with distance until it vanishes, or becomes infinitesimal, on the horizon.

It should be pointed out that size perspective and linear perspective, when considered as retinal gradients, are stimulus-correlates of continuous physical distance. They are to be distinguished from the traditional "cue" for distance-perception of the apparent size of familiar objects, i.e., of objects whose real size is remembered from past experience. The comparison of an absolute retinal size and a remembered size and the inferring of the distance, assuming it to occur, is not an adequate explanation for the perception of a continuum of distance. The explanation proposed here does not assume the perceiving of absolute sizes as such but only the ability to react to a continuous gradient of retinal sizes.

The facts of texture perspective and size perspective as described refer to the *retinal image* of the terrain in two dimensions. The resulting *perception* of an extended terrain in three dimensions is characterized by objects and terrain features which do *not* shrink in size toward the horizon. Instead, they appear to maintain a substantially constant size and are perceived at a distance. The relation between this constancy of perceived size and the perception of distance will be discussed later.

The Retinal Gradient of Velocity During Movement of the Observer

A third stimulus variable for the perception of distance is one which is particularly applicable to the flying situation since it occurs during movement of the observer. It bears some relation to the cue of monocular motion parallax. When an observer moves, and particularly when he is flying or driving, the visual world is represented by images which also move across the retina of the eye. The simplest form of this retinal motion may be described by the statement that the image of the world expands radially outward on the retina as one moves straight forward. The expanding optical picture ahead as one drives a car is the most familiar example, and it has probably been noticed by nearly everyone. If, instead, one looks backward, the world (considered as a flat image) contracts inward on the retina as one moves away from it. The center of this expansion, the point from which it radiates, is that point toward which the observer is moving. There is a center of contraction at the opposite pole, i. e., the point he is moving away from. During ordinary locomotion, the center of expansion is on the horizon.

Now under such circumstances the retinal motion of the image corresponding to the terrain is subject to the principle of perspective. There exists, in other words, still another type which will be called retinal motion perspective. The *rate* of expansion of the image of any point or object is inversely proportional to the distance of that point or object from the observer. There is, in other words, a continuous gradient of the velocity of the ground as it "goes by;" the gradient begins with a maximum at the points of the terrain nearest the observer and ends with zero movement at the horizon. This rule holds no matter in what direction one looks. Such a gradient of velocities is capable of determining a continuum of distance and, within this dimension, the distance of any point or object is determinate from its retinal velocity.

When a retinal gradient of velocity exists in the way described, the perception which results is not that of a visual environment which moves but of a stationary world in which the observer himself moves. If the observer is not moving but is, let us say, sitting at a desk, it is nevertheless true that his head will move from

time to time and that the image of his visual world moves on the retina. Optically speaking, the world is "alive" with retinal motion produced by only the ordinary slight displacements of the head and body, and the gradients of motion which result are ever present stimuli for the visible continuum of distance.

The description above leaves out of account a number of the characteristics of motion perspective, and makes no mention of several complicating factors. When the motion of the observer is not parallel to the terrain, as when a pilot lands an airplane, the formulation given must be modified. The effect of eye movements on motion perspective also needs to be considered. These matters will be discussed in a later section. For the present purpose of listing the sensory bases for distance perception, the description above will suffice.

The Retinal Gradients Arising from Atmospheric Transmission of Light

The cue of aerial perspective as ordinarily described provides another kind of retinal gradient which is a continuous correlate of distance. The retinal image of a terrain stretching away to the horizon is constituted by light which at one extreme has passed through only a few feet of air and at the other extreme has passed through many miles of air. The character of the light stimulus varies with the amount of atmosphere through which it has been transmitted. The resulting color quality becomes less saturated and bluer with increasing atmospheric distance. The color is also described as being increasingly blurred or film-like in appearance with increasing lengths of aerial transmission, and the outlines within the image become less sharp. It is possible that these latter variations should be considered in relation to the texture variable. The exact stimulus-variations involved have not been worked out in detail. They are effectively employed by painters but they have not been fully described in terms of physiological optics.

The Retinal Gradient of Binocular Disparity

A number of visual stimulus dimensions have just been defined which are concomitants of distance and which are presumably stimuli for the perception of space as the flier sees it. They are all based on gradients of stimulation in a single retina; that is to say, they do not depend on differences in stimulation between the two eyes. There is, in addition, however, the fact of binocular retinal disparity, or stereopsis, which has received most of the attention devoted to the problem of distance perception in the past. This variable can be defined, like the others, in terms of a gradient of stimulation, with only the addition of the fact that the stimulation referred to is a binocular rather than a monocular effect.

Assuming for the moment that the observer's eyes are fixated on the horizon, the retinal image of the terrain in the right eye will differ from that in the left eye. This difference at any given point is called retinal disparity, and is due to the different positions of the two eyes in relation to the terrain. Near points and objects on the terrain are displaced horizontally in the image of one eye relative to the other. This relative displacement decreases with increasing distance and becomes zero at the horizon itself. There is, in short, a gradient of disparity in the combined retinal field. It is, like the others already described, a vertical gradient, running up the field from the observer's body at one extreme to the horizon at the other. Any point on the terrain corresponds to a disparity which is inversely proportional to the distance of that point from the observer. It must be supposed that this variable is a stimulus-correlate of perceived distance.

This description holds true when the eyes are fixated on the horizon. If instead, the eyes are fixated on a near point, the disparity is zero at that point and reaches a maximum at the horizon. But this disparity is opposite in kind to that existing in the former situation; it is "uncrossed" rather than "crossed," or positive where the former was negative. The *gradient* of disparity with respect to its sign is therefore the same when the eyes are fixated on a near point as when they are fixated on a far point, or for that matter when they are fixated at any point. An increase in positive disparity being equivalent to a decrease in negative disparity, the gradient may run from minus to zero or from zero to plus and still be the same gradient. The stimulus which is concomitant with distance, therefore, is not simply disparity as such but disparity relative to a gradient which may lie anywhere on a scale of negative to positive.

The Relation of Other So-Called Cues for Depth to the Variables Above

All of the traditional cues for depth perception have been incorporated or reinterpreted in the variables listed, except for interception or superposition of contours and the distribution of shadows and shading. Interception is capable of determining the relative distance of two or more objects but, by its very nature, it is not a variable which can establish a continuum of distance. It has to do with the establishing of the figure-ground relationship and the relation of "behind" or "in front of" rather than with distance perception as such. The distribution of shadows produced by objects and the gradients of shading appearing on three-dimensional shapes are determiners of what is properly called relief or relative depth, but this is not the same thing as the sensory continuum of distance. They will not be discussed further, nor

will any analysis be given of the retinal gradient associated with accommodation—a kind of "blur" gradient.

Methods of Reproducing the Stimulus Variables for Distance

In order to construct tests or carry out research on the perception of distance, the stimuli must be controlled and systematically varied. Of the variables listed, only that of retinal disparity, together with accommodation and convergence, has been systematically utilized in experiments or tests. The variables of texture, size perspective, linear perspective and aerial perspective are, within limits, capable of being reproduced by photographic and pictorial means but this has not been done by psychologists. A great deal of attention has been paid to these factors by painters and photographers but only for their own limited purposes. Retinal motion perspective has received no study whatever, possibly because a motion-picture technique is the only feasible way to reproduce it, and because the use of motion pictures for experimental purposes is still almost wholly undeveloped.

Tests for aerial distance perception are limited to a small number of possibilities. The subjects may be taken into the open air and presented with a real spatial situation. All the sensory variables are then present if the situation has been properly selected or arranged. But this method is hardly feasible for group testing on a large scale. Or the subjects may be presented with still photographs or motion pictures of a selected or arranged spatial situation. If the limitations imposed by the method are complied with, all the sensory variables except the binocular ones can be represented. Once the nature of the variables is understood, they can be specified both in the spatial situation and in the photographic reproduction. The variable of retinal disparity could be included in the test situation by presenting stereoscopic photographs. Individual stereoscopes would, however, be required and the subjects would have to be trained to use them. There is a third alternative, which is the one adopted in nearly all the existing tests, namely the setting up of a "room-sized" spatial situation by means of an apparatus. But such a situation does not represent aerial space. All the tests of this type, moreover, have excluded the monocular sensory variables on the grounds that they were not intrinsic or innate factors in depth perception.

The method adopted for reproducing the distance variables in the research to be reported was the second of these alternatives. Still and motion picture photographs were employed, both of artificial and of natural situations, under controlled conditions. Evidence will be presented that the photographic representations actually do reproduce the dimension of distance, and that they make possible discriminations only somewhat less accurate than

those which are made in equivalent "real" situations in the open air.

THE PROBLEM OF TESTING—TYPES OF JUDGMENT INDICATIVE OF THE ABILITY TO PERCEIVE DISTANCE

Any perceptual test or psychophysical experiment involves judgments or discriminations scorable as correct or incorrect. Although distance is perceived as a continuous dimension, judgments of distance must be made with respect to specific objects at specific distances. The kind of distance perception required for flying is not easily represented by judgments which can be set up under controlled conditions. Several kinds of judgments are possible for use in devising tests. They will be described and reasons given for rejecting some in favor of others.

Judgments of the Relative Distance of Two Objects Side by Side

The familiar pin test and its variants, including the Army Air Forces form named the Howard-Dolman Test, require the subject to discriminate the relative distance of two pins or sticks. The pins are seen through a window which excludes any view of their ends and of the surrounding environment. The distance of the pins from the subject in the AAF test is 20 feet. No continuous gradient of stimulation is present in the situation. It is at least a question whether such a judgment of relative distance is analogous to the judgments required in flying. Other criticisms of the test have already been made on the grounds that it reproduces mainly the cue of binocular parallax and that the absolute distance involved is limited to "room-size" space.

Tests of stereoscopic acuity, employing either stereoscopes or a polarized light-and-goggles system, are subject to the same objections in a more acute form. All of them require judgments of the relative distance of a standard and a comparison object. This type of judgment can be used only when the two objects are seen in "empty" space, i. e., one without a visible substratum extending from the observer to the objects. If a floor or ground is visible, the point of contact of the objects with the ground "gives away" the judgment of relative distance. The point of contact with the ground which is uppermost in the field, considered as a picture, determines which object is behind the other. The relative distance of two objects on the ground side by side is so easy to discriminate that it becomes valueless for a test of distance perception.

Judgments of the Absolute Distance of a Single Object

It would be desirable to set up a situation in which the observer had to judge or estimate the absolute distance of a test object in terms of an arbitrary scale. The only scale which appears to be

usefui, however, is the conventional scale of yards, feet, or meters. Efforts to utilize this type of judgment were made in the early stages of this study, but the judgments proved to be highly unreliable. The average observer does not have a consistent idea of one hundred or two hundred or even ten yards. The method seemed to be impracticable for testing, although if time were available to train the observers it might prove valuable for research purposes.

In an open-air situation it is possible to require the observer to judge the distance of an object relative to the distance of another object, or series of objects, *in a different direction on the ground*, and therefore not seen in the same field of view. Since the objects discriminated are not side by side the cue described above would not give the judgments away. A study was carried out by this method in the early stages of this investigation and it proved to be successful. For a number of reasons, however, this situation cannot be reproduced by photographs without introducing artifactual cues to distance, and the method was therefore given up as a possibility for use in group testing.

Judgments of the Size of a Far Object in Relation to That of a Near Object (Size-Constancy Method)

It is much easier to set up comparisons of sizes than comparisons of distances. If a subject is able to judge the true size of a distant unfamiliar object, he does so only because he sees the true distance of the object. Its true size cannot be judged from the retinal size of its image alone; it must be judged in relation to a gradient of sizes (and other stimuli) which constitute its distance. The ability to estimate the sizes of distant objects or, specifically, to match them accurately with the corresponding sizes of near-by objects is therefore indicative of the ability to estimate their distance.

Judgments of size-at-a-distance, employing arbitrary test objects of unfamiliar size, can readily be set up in both artificial and natural environments, and the situations can be represented by photographs. A considerable amount of preliminary research was carried out on both this type of judgment and on the spatial situations in which it can be required. If the size of the distant test object, for example, is given in terms of linear perspective (as the width of railroad tracks is given) the judgment becomes too easy to be of value. If this and other artifactual cues are eliminated, however, judgments of the true size of distant test objects can be shown to be practicable and reliable in distance perception research. This type of judgments was therefore adopted for the construction of a photographic test of distance perception (CP212A) a report of which is given in the next section.

Judgments of Absolute Distance During Locomotion

A fourth kind of distance judgment is possible which would appear to be even more similar to the judgments demanded in flying than those listed. By means of motion picture presentation the view of the terrain during the landing glide of an airplane can be reproduced on a screen. All the stimuli for distance perception except binocular stimuli are present. A judgment of distance from the ground ahead can be required in terms of a temporal scale. The subject need not be asked to estimate altitude in feet; he can be shown a trial run during which a certain "critical" altitude is designated by the voice accompanying the film. In a subsequent run the critical distance must be judged in terms of a series of temporal intervals also designated by the voice.

A motion picture test of this type was planned, but not completed. Inasmuch as judgments of the *direction* of the glide in such a scene are more easily scored and are probably as important as judgments of distance, the former were utilized in a test which was pushed to completion, the Landing Judgment Test CP505E. The background of this test will be described in a later section.

A PHOTOGRAPHIC TEST FOR DISTANCE PERCEPTION

The Relation Between Size Constancy and Distance Perception

A problem which has not been considered in sufficient detail is the relationship between size constancy and the perception of distance. Since the most practical method of constructing a distance perception test is based upon the use of photographs and the method employed in experiments on size constancy, it becomes very important to point out the relationship between size constancy and distance perception. For example, it might be argued that there is little connection between judgments of the real size of objects and perception of the continuous space between the observer and object. If so, there is, of course, no justification for basing a distance perception test upon the accuracy with which the sizes of objects at a distance are judged.

Size constancy is the term used in psychology to refer to the fact that the seen size of objects does not diminish as they become more distant, as do the retinal images corresponding to these objects. Instead, objects appear to maintain an approximately constant size. The limitations and explanation of this fact have been the subject of a great deal of experimental research in the psychological laboratory, although very little of this research has been carried out at large distances in the open air. The experiments consist essentially of comparing the size of a far object with the size of a near object.

One may conclude from these experiments that to speak of the visible size of an unfamiliar object is meaningless apart from its visible distance. If the distance for some reason is not perceived, the size becomes indeterminate. The size of the moon in the sky is a good example. Its size is vague and indefinite, appearing to one individual to be that of a dime, to another that of a dinner plate, and to still another it may appear as obviously miles in diameter. The explanation is that the *distance* of the moon is indefinite, since there are no gradients of retinal stimulation in the sky to produce a definite impression of distance.

The stimulus for the perception of the size of an object is therefore not simply the size of the retinal image; it is this in combination with the stimuli for distance. Both retinal size and distance have to be sensed in order to perceive real size. A perception of size therefore always involves a sensed distance.

The way in which retinal size and distance combine physiologically to give a reasonably constant perception of objective size is not understood. It need not be discussed in this chapter, except to suggest that the physiological process corresponding to seen distance is reciprocal to the physiological process corresponding to retinal size. More of one process would compensate for less of the other, and hence a small distance and a large retinal image would yield the same perceived size as a large distance and a small retinal image.

There is ample experimental evidence that size constancy is brought about by distance perception. The results of these experiments are usually given in terms of the amount of constancy which the size estimates show, varying from zero constancy if the size is perceived in accordance with the retinal image to 100% if the size is perceived objectively. The amount of constancy is usually high in such experiments, but it is lowered if the conditions for viewing the distance are poor. For example, size constancy is much reduced if the objects are seen monocularly in very dim illumination. Presumably the stimuli for distance are weakened in the dark. In viewing luminous objects in total darkness, when all stimulus gradients for distance are eliminated, the situation would be like that of the moon in the sky, and there should be no size constancy whatever.[7]

Constancy of Size Perception vs. *Accuracy of Size Perception.* The method of judging size-at-a-distance employed in the constancy experiments may be used for the practical purpose of testing distance perception. But the measure of the amount of constancy employed in these experiments (the index of constancy, or percent of constancy from zero to 100) is not a useful one for a test

---

[7] Holway, A. H. and Boring, E. G. Determinants of apparent visual size with distance variant. *Amer. J. Psychol.*, 1941, *54*, 21-37.

of distance perception. What is needed for such a test is a measure of the objectivity or accuracy of the judgments of size-at-a-distance. Individuals may overestimate as well as underestimate the size of a distant object. In either case there is, from a practical point of view, an error in perception. But from the theoretical point of view an overestimation represents more than perfect constancy, and the index is over 100. The effect is sometimes called "over-constancy," and it is troublesome to the theorists. Since overconstancy obviously does not produce a better or more objective perception than underconstancy, this measure is inappropriate for a test. Accuracy of size perception, on the other hand, is both appropriate and simpler to compute. It is simply the closeness of the subject's estimate of the size of a distant object to its actual size, irrespective of the direction of the error.

An Experiment to Validate the Photographic Method of Representing Distance

Before a photographic test of distance perception can be constructed it is necessary to "validate" the photographic method of representing space. The question to be answered is the extent to which tridimensional space can be duplicated in a photograph which is itself a flat surface. It is implied by the theory of retinal gradients already stated that photographs can represent distance in a properly chosen scene much more adequately than is usually supposed; and conceivably can portray the distance about as well as the eye can see it. This hypothesis, upon which a photographic test for distance estimation is based, requires experimental verification.

The research to be reported here is a study of the nature of photographic space as well as being one step leading to the construction of a psychological test. There has apparently been no comparable experimental attempt to determine quantitatively the power of single photographs to portray tridimensional space, and this seems to be the first investigation of the subject. Ames[8] has studied the portrayal of depth in paintings. By means of a cylindrical lens he reduced the surface quality of the canvas and enhanced the depth effect. In this way he was able to find errors in the paintings and analyze certain factors of technique producing depth. Schlosberg, as already mentioned[9], has also demonstrated that the depth represented in photographs could be made visible by using optical devices. Neither of these studies, however, has determined, in an experimental and quantitative manner, the effectiveness of photographically represented distance in comparison to that of real distance.

[8]Ames, A., Jr. Depth in pictorial art. *The Art Bulletin*, 1925, *8*, 5-24.

[9]Schlosberg, H. Stereoscopic depth from single pictures. *Amer. J. Psychol.*, 1941, *54*, 601-605.

The experiment was set up to answer the following questions: With what degree of accuracy can subjects judge the size of unfamiliar objects, never seen at close range, which are shown him at various distances up to half a mile on a terrain carefully selected as being favorable for judging distance? With the subjects still kept in ignorance of the objects and their success in judging them, how accurately can they make the same judgments (in a different order) from photographs of the same terrain and the same objects? The significant difference between the two tests was the difference between seeing the situation directly and seeing it photographically.

The judgments of size-at-a-distance were made in terms of a standard scale of sizes in the foreground, the distant object being matched with the standard object judged to be of the same size. The situation is shown in Figure 9.5, which represents one of the test objects (75 inches high) at one of the distances (112 yards). Precautions were taken to prevent the subjects from assuming that the distant objects were probably about as high as the middle objects of the scale. They were told that the distant objects might be off the scale in height and that they should so estimate whenever the object so appeared. In order to reinforce this attitude, shorter and taller objects were actually presented at irregular intervals. It can be assumed that the subjects knew or inferred nothing about the size of the objects except for what they could see immediately as a result of the perceived distance.

The Experiment

Subjects. A group of 15 subjects was used in this research, consisting of 14 enlisted male personnel and 1 female at Santa Ana Army Air Base. Thirteen of the same subjects were observers in the photographic duplication of the outdoor experiment. All results given refer to the 13 subjects common to both experiments, except when specifically noted.

The Spatial Situation. In order to obtain judgments of the size of objects in a realistic situation, it was finally decided after considerable experimentation that a natural scene would provide better tridimensionality than any of a number of types of artificial scenes. It was necessary, therefore, to go outdoors for the research. There had to be found a long stretch of level ground which possessed adequate stimuli for distance perception. After considerable exploration and experimental photography, a large open field adjoining a river bed near Santa Ana Army Air Base was selected as satisfactory for the experiment. The field consisted of cultivated land; it was almost perfectly flat and level and had a fairly coarse but evenly-structured surface providing good "texture-perspective" for distance. The field was bounded on the left

by brush growing on the river bank and on the right by rolling hills of cultivated ground. These boundaries merged in the far background to give some appearance of size perspective to the scene. It is shown in Figure 9.5. The linear perspective produced by regular furrows was avoided since such furrows would provide a cue by which size could be estimated at any distance. An object, for example, might be estimated as two and one-half furrows high. Here the objects whose size (in this case height) was to be judged were narrow white stakes which could easily be stuck into the sandy ground. The *distant* stake, called the test object, was presented in one of six different sizes at each of six distances. The size of the test object was to be estimated in terms of a standard scale of sizes in the near foreground. For this purpose, 15 stakes were set into the ground in an arc at a distance of 14 yards from the observer and were numbered from 1 to 15, starting with the shortest. The stakes were equi-distant from each other (20 inches apart) except for a large gap of eight feet in the center of the arc. Through this gap the subject could see the distant stake whose height was to be judged in terms of the scale. Number 1 of the scale of standards was 27 inches high and the stakes increased in size by 4 inch steps up to number 15, which was 83 inches high. The width of these stakes (and also the width of the test objects) varied in an irregular manner between about 2 inches and 4 inches.

The six distances at which judgments had to be made were 28, 56, 112, 224, 448, and 784 yards. If the distance from the observer to the standard scale be called "X," then the distances of the test objects form the series 2X, 4X, 16X, 32X, and 56X. The field was not sufficiently long to permit use of a distance of 64X, or 896 yards. The objects at 784 yards, however, were beginning to be difficult to see since the corresponding retinal images were very small.

Four different sizes of the test object were presented, with the addition of a very short and a very long stake. They were 15, 63, 67, 71, 75, and 99 inches high. Four of these stakes were matched in height with the standards numbered 10, 11, 12, and 13. These were the test objects of the experiment. Their mean height was 69 inches. The short and the tall test objects were introduced into the experiment in order to guarantee that the subject not assume that he must restrict his judgments within the limits set by the ends of the standard scale. Subjects were therefore free to make any size judgment whatever which seemed to them correct. Judgments of these stakes were recorded but were not used. The shortest object was 12 inches shorter than the shortest member of the scale and the tall stake was 16 inches taller than the tallest stake of the scale.

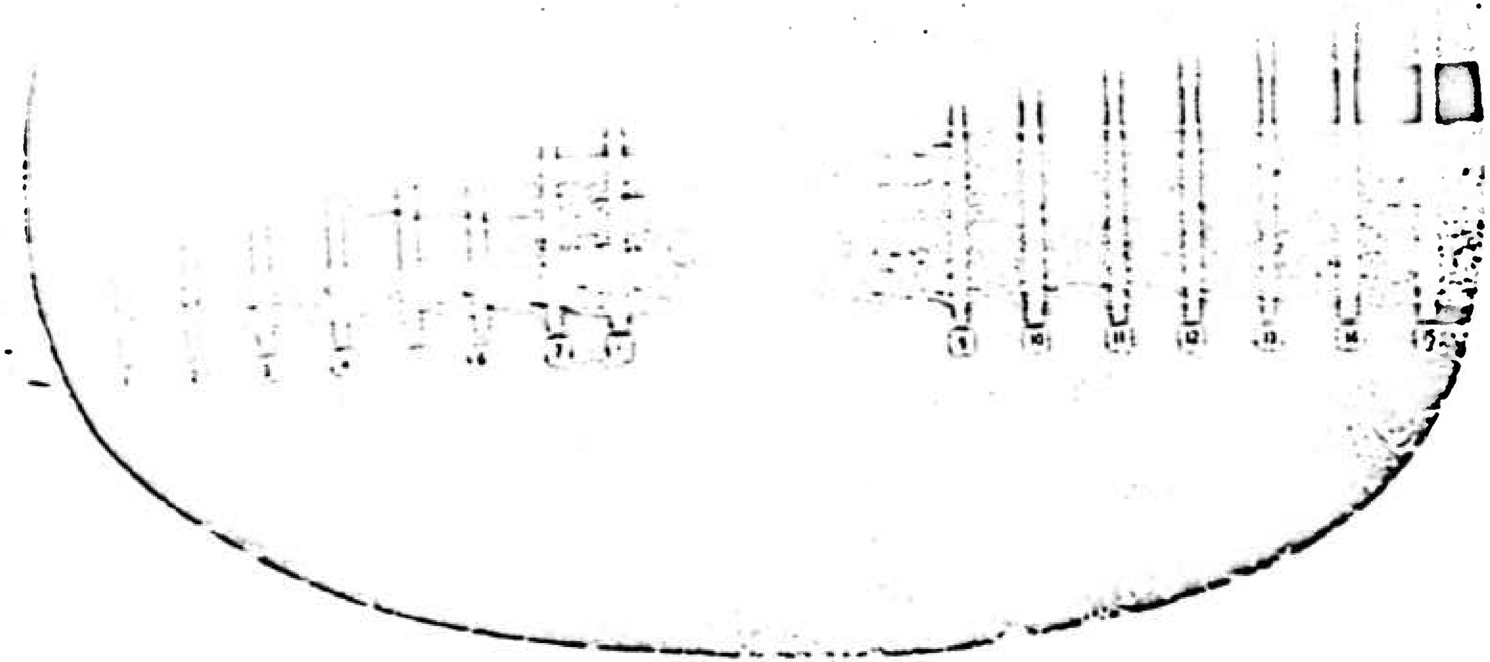


FIGURE 9.5.—Judgment of Size-at-a-Distance.

Procedure. In both the open-air experiment and the photographic repetition of the experiment, the subjects were required to match the height of the test object in the distance with the height of one of 15 standards in the foreground. Each estimate was recorded on an answer sheet as a number from 1 to 15. Although as much time as was needed for their estimates was given to the subjects, they were urged not to delay making the judgments. The time required did not exceed about 10 seconds.

The Instructions. The instructions given to the subject were as follows:

This experiment is intended to test your ability to judge the correct height of objects when they are seen at a distance. You will be required to estimate the real height of a white stake out there in the field in terms of the numbered series of 15 stakes directly in front of you. Choose the height from 1 to 15 which corresponds to the height of the stake in the distance. Note that the width of the stakes varies in a random fashion; hence you must be careful to judge height alone and disregard width.

You will make a series of 25 judgments at each of six distances, starting at the greatest distance. Record each estimate in order, writing down the height, i. e., the number from 1 to 15, on the cards you have been given. At the end of each 25 judgments, you will begin to use a new card. Note that the cards contain spaces from 1 to 25 and that you have 6 cards. At the larger distances, some of the judgments may be relatively difficult. Record an estimate for each trial even though you may not be completely sure of it.

The stakes you will see at the various distances may vary from an extreme below the smallest height on the scale of 15 to one higher than the tallest on the scale. If you judge the height of the distant stake to be smaller than any of the 15, record "S" for that trial; if you judge it to be higher, record "L." If, however, it falls anywhere within the 15, make the most accurate judgment you can. Make no assumption about the different heights presented at any one distance, since the stakes presented may vary from one distance to another. Consider each judgment very carefully and make each one the best estimate of which you are capable.

Stand within this circle marked on the ground when you make your judgments. As soon as you have recorded your estimate, *face in the opposite direction until the signal is given for the next trial.* Above all, do not speak, gesture, or even look at your neighbors while making your judgment, and do not under any circumstances compare notes between trials. Your estimate must be completely independent and unbiased.

You will first be given five practice trials in which the stake to be judged is at the same distance as the 15 standards. Record your estimates on the card marked "practice." These judgments will probably be very easy.

The Outdoor Experiment. After the subjects were given their instructions, five trials were run off for practice with the test objects at the same distance as the standards, i. e., 14 yards. A random series of 25 trials at each of the six distances was then presented, starting at the greatest distance and ending with 25 trials at the same distance as the standards. One hundred and fifty judgments were thus obtained from each subject. The sub-

jects did not know how many stakes there were in the set of test objects, since they did not see them before the beginning of the experiment. They did not see the stakes being set up between trials. They did not know, except for what they could see during the trials, either the real size of the stakes or the distance at which they were presented.

The Photographic Experiment. Photographs were taken on a separate occasion of the scene described above, one photograph for each trial, with the camera at the same position occupied by the subjects in the original experiment. In effect, the photographic experiment was an exact repetition of the outdoor experiment except for the stimulus situation presented. Photographs were not taken, however, of the test objects at the greatest distance, 784 yards, since it had been determined that the images of the stakes at this distance would have been too small to be adequately represented on a photographic print. Six photographs, one for each test object, were made at 14, 28, 56, 112, 224, and 448 yards, making a total of 36 photographs available for the experiment. The photographs were taken with a 4 x 5 Speed Graphic (focal length 5¼ inches). The negatives were later enlarged to make 8 x 10-inch glossy prints. Every effort was made to procure maximal and natural depth portrayal in the photographs. When presented to the subjects, an oval frame was placed over the rectangular print, using a paper mask for this purpose. Uniform illumination of the photographs was provided in order to eliminate surface reflections from the glossy paper. The subjects were seated at such a distance from the picture that the retinal image would be identical in angular size to that in the outdoor situation, except that the field of view was of course cut off by the frame of the photograph. The pictures were viewed, in other words, at the so-called "optimal viewing distance." The procedure of the experiment was as similar as possible to that employed in the outdoor experiment. As in the latter experiment, they were given five trials for practice with the test objects at the same distance as the standards, followed by 25 trials at each of the six distances, beginning with the farthest distance, which in this case was 448 yards, and ending with 14 yards.

The Results

Accuracy of Size Judgments in Real Space and in Photographs. The accuracy with which the sizes of the four test objects were judged at various distances in real space and photographs is given by the average errors of the judged sizes from the true sizes of the test objects. The average errors were computed for each test object at each distance by obtaining the sum of the difference in inches between each judgment and the correct size and dividing

by the total number of judgments of that object. "L" judgments were scored arbitrarily as two steps (8 inches) beyond the end of the scale. They were few in number. There were no "S" judgments. The average errors for the real space experiment are given in table 9.1-A, while those for the photographic experiment are found in table 9.1-B.

TABLE 9.1.—*Average error in inches from the true size of objects judged at different distances in real space and with photographs*

A. OUTDOOR SPATIAL SITUATION (15 SUBJECTS; 100 JUDGMENTS)

	63 inches	*67 inches*	*71 inches*	*75 inches*	*Mean of average errors*
Same distance (14 yards)	.48	.76	.90	1.48	.90
28 yards	3.84	4.24	4.80	4.40	4.32
56 yards	3.80	4.04	4.80	5.32	4.49
112 yards	5.92	6.16	5.88	6.32	6.07
224 yards	8.84	7.80	7.72	6.52	7.72
448 yards	9.00	9.44	8.04	7.60	8.52
784 yards	8.32	10.64	9.36	7.76	9.02

B. PHOTOGRAPHIC SITUATION (13 SUBJECTS; 65 JUDGMENTS)

	63 inches	*67 inches*	*71 inches*	*75 inches*	*Mean of average errors*
Same distance (14 yards)	.37	1.11	1.11	1.11	.92
28 yards	7.38	6.52	7.63	7.38	7.22
56 yards	7.69	8.86	8.43	8.18	8.29
112 yards	8.25	8.18	7.51	7.51	7.86
224 yards	6.40	7.81	7.14	7.32	7.17
448 yards	8.37	10.03	9.97	10.34	9.68

The most striking fact shown in these tables is that the average error in estimating the size of distant unfamiliar objects is small. Even at a distance of 784 yards, when the objects, averaging 69 inches in height, are so far away that they are beginning to be difficult to see at all, the average error is only 9 inches. At this distance, the retinal image of the object is one-fifty-sixth of the size it possesses at 14 yards. It must be concluded that the distance itself is perceived with some precision. Moreover the errors made in photographic space are not much larger than those made in the real situation. Between the distances of 28 to 448 yards, the range of the average errors for real space is from 4.32 to 8.52 inches and for photographs from 7.17 to 9.68 inches. In both situations the average error did not far exceed two steps of the standard scale on either side of the size which was the correct match even at a distance of a quarter of a mile. Size perception is evidently determinate up to very great distances, much greater than previous experiments have indicated, and the cues underlying this perception can evidently be duplicated by photographs much more effectively than the binocular theory of distance perception would lead one to suppose.

Comparison of the means of the average errors (i. e. of the mean test object, 69 inches high) shows that there is a progressive decrease in the accuracy of size judgments with each increase of distance in real space. This result can be expected from the fact that distance enters into the judgments of size, so that as distance

increases judgment becomes more difficult. The trend of decreasing accuracy with distance is not as regular for the discriminations made in the photographs, however, as may be seen in table 9.1–B.

A significant fact is that when the distance of the standard and comparison objects is the same (14 yards) there is only a small amount of error, on the average less than one inch in both real and photographic space. However, with the test object placed at 28 yards from the observer, or 14 yards beyond the standards, there is a marked increase in mean average errors. In both real space and photographs the decrease in accuracy is greatest from 14 to 28 yards. It is obvious that sizes can be discriminated most accurately when the distances are identical. In this case, size discriminations are independent of distance discriminations.

Average Size Judgments and Their Variability in Real Space and Photographs. The estimates of all subjects were converted into inches and the results for all four test objects have been combined into one theoretical test object 69 inches in height. The mean size estimate and the standard deviation of the estimates are given for all distances in table 9.2. The standard deviation of

TABLE 9.2.—*Means and standard deviations in inches of size judgments made at different distances in real space and photographs*

[All four test objects combined; mean size of test object = 69 inches]

Distance in yards of test object	*Real space: Subjects-15; N = 300*		*Photographs: Subjects-13; N = 260*	
	Mean	*S.D.*	*Mean*	*S.D.*
14 (same)	69.67	2.11	69.98	1.79
28	68.47	5.29	73.15	7.36
56	70.80	5.83	74.93	7.60
112	73.57	6.27	72.85	8.15
224	73.62	7.67	70.58	8.43
448	72.44	9.37	65.40	10.72
784	72.28	10.18	......	......

the estimates represents a measure of the variability of the judgments, and the mean estimate represents the actual magnitude of th size perceptions at various distances or, in other words, the *trend* of the errors. The mean size estimates also give evidence of the degree to which size constancy has been attained in the perceptions.

It can be noted that in both real space and photographs the variability of the size estimates rises when the standard and variable are at different distances, and increases with each increase of distance. With respect to the relationship between variability of size judgments and increase in distance both spatial situations produce the same effect. The variability at each distance, however, is somewhat greater in photographs than in real space. These results are consistent with the hypothesis that errors in size perception at a distance reflect errors in visual distance perception; but that the accuracy of such perception is reduced in photographs.

Of equal significance to the validation of the photographic method of representing distance is the fact that, on the average, the test object is seen at nearly its true size at distances as far away as a quarter of a mile. This is true of the size estimates made with photographs as well as those in the outdoor situation. In fact, in the outdoor experiment there is no evidence whatever of a falling off of perceptual constancy with an increase in distance up to almost one-half mile. In the size estimates made from photographs a small decrease in constancy occurred at the distance of 448 yards. However, in general the reverse is true; the size tends to be overestimated. The individual observers varied in this respect, some underestimating, but the majority overestimating size at a distance. These facts are of theoretical importance since they are not in accordance with predictions made by some investigators of size constancy in the laboratory. The point to be stressed here is that the maintenance of size constancy at far distances in real space and photographs means that distance is clearly discriminable at long ranges in both situations.

Comparison of the Consistency of Size Judgments of the Same Objects Seen at Different Distances in Real and Photographic Space. The question arises whether each observer's size judgment of an object will be similar from one distance to another. Will individual A tend to overestimate sizes at all distances, and B underestimate them? The consistency of the size estimates of the observers when observing at different distances was obtained by the method of correlation. In both the outdoor and photographic experiments the observer's mean size estimate for each test object at each distance was ranked on the basis of its magnitude. This resulted in a rank order of mean size estimates based upon magnitude for four test objects (63, 67, 71, and 75 inches) at each of four distances(56, 112, 224, and 448 yards), or a total of 16 ranked orders of mean size estimates. Rank-order correlations were then made between the magnitude of size estimates at 56 and 112 yards, 112 and 224 yards, and 224 and 448 yards. The correlations for the outdoor experiment are given in table 9.3, and those for photographs are in table 9.4.

TABLE 9.3.—*Rank-order correlations between the mean size estimations of 15 observers for each test object at various distances in real space*

[Five judgments obtained at each distance]

Size of object	*56 and 112 yards*	*112 and 224 yards*	*224 and 448 yards*
63 inches	.88	.90	.82
67 inches	.57	.60	.73
71 inches	.49	.62	.63
75 inches	.47	.60	.84

Mean r is .65.
Range is from .47 to .90.

TABLE 9.4.—*Rank-order correlations between the mean size estimations of 13 observers for each test object at various distances represented in photographs*

[Five judgments obtained at each distance]

Size of object	*56 and 112 yards*	*112 and 224 yards*	*224 and 448 yards*
63 inches	.69	.85	.50
67 inches	.82	.92	.67
71 inches	.84	.90	.52
75 inches	.79	.88	.49

Mean r is .73.
Range is from .49 to .92.

The average size judgments in both situations are fairly consistent for each test object and distance. In the outdoor experiment the range of correlations is from .47 to .90 with the average at .65; in photographs the correlations range from .49 to .92 and the average is .73. These results imply that there are consistent individual differences in the judgment of size at a distance, and that these individual differences manifest themselves in photographic as well as in real space. Observers tend to maintain their constant errors in judging size despite variation of distance. Since a photographic test of distance perception must utilize a number of distances these results encourage the belief that the test developed as a result of this experiment will be reliable.

The individual differences among the observers in their average size judgments had a considerable range. The judgments of some observers represented overestimates, and of others, underestimates; some were accurate and others inaccurate. For example the mean perceived size of the 63-inch object at 224 yards ranged among different observers, from 60 to 81 inches in the outdoor situation and from 57 to 81 inches in the photographic situation. These individual differences also encourage the belief that a test of ability to perceive distance accurately is feasible and necessary.

Correlations Between the Magnitude of Size Judgments of Objects in Real Space and the Magnitude of Size Judgments in Photographs. The question also arises whether an observer's size estimates of the objects will be similar in the photographic situation to what they were in the outdoor situation. If individual A judges the test objects to be higher than do the others in the group, will he continue to do so when presented with photographs, and if individual B underestimates relative to the group, will this tendency persist in the photographic situation? If the photographs used in this experiment adequately portray the real situation, there should be a consistent relationship between the individual differences in the magnitude of size-distance estimates manifested in the two situations. This result would tend to "validate" the photographic method.

The correlations are presented in table 9.5. They were obtained in the following way. An observer's mean size estimates for

Table 9.5.—*Rank-order correlations between the magnitude of size judgments of objects in real space and in the magnitude of size judgments of the same objects represented in photographs*

[Based upon 25 judgments obtained at each distance from 13 observers]

Distances in yards	*Test object size*			
	63 inches	*67 inches*	*71 inches*	*75 inches*
56	.88	.88	.85	.87
112	.77	.47	.35	.84
224	.55	.87	.88	.82
448	.80	.72	.68	.83

Mean r is .73.
Range is from .35 to .87.

each test object at each distance were ranked on the basis of their magnitude. Sixteen rank-orders of mean size estimates were thus obtained for four test objects at each of four distances (56, 112, 224, and 448 yards). This was done for both the outdoor and photographic experiment. Correlations were then calculated between the rank order of estimates of equivalent objects and distances in real space and in photographs.

The average correlation between the magnitude of size judgment in the two situations is .73, and the correlations range from .35 to .87. The correlations are, in general, sufficiently high to support the statement that the size-distance judgments made in photographs are similar in magnitude to those which occur outdoors. Since what is correlated is the magnitude of the observer's average judgment, a further implication may be pointed out. Constant errors in estimating size at a distance are characteristic of the individual, since individual differences are consistent in the two situations being compared. Observers who overestimate size of objects in real space are likely to have relatively high estimates of the size of the same object at the same distance in photographs. Observers who underestimate size in real space also tend to underestimate in photographs.

Implications of these Results for the Theory of Aerial Space Perception and the Theory of Perceptual Constancy

The results of the foregoing experiment imply that a novel and, on first impression, inappropriate method may be used to study and test the ability to judge distance—the method of photographs. They also lead to a number of novel conclusions regarding the perception of space at large distances and the theory of size constancy. It has been rather generally assumed on the basis of laboratory results that constancy of perceived size falls off with increasing distance, or at least that apparent size eventually decreases at very great distances. If this were true it might mean, according to the reciprocal theory presented, that distance itself is shortened at long ranges and that there is a constant error present in the perception of aerial space as distinct from local or "room-sized" space, tending to make great distances in general underestimated.

In the experiment reported, size perception was studied up to the point where the size showed signs of beginning to disapear optically. The question is, do sizes become smaller in *perception* before they reach their "vanishing point." Does the apparent size of objects ever decrease at large distances?

Table 9.6 and figure 9.6 show the apparent sizes of the four test objects in inches at the various distances in the outdoor situation. The perceived sizes at 784 yards have not been plotted in

TABLE 9.6.—*Means and sigmas of size estimations for each test object when presented at different distances in real space*

[Based upon 5 judgments of each object at each distance obtained from each of 15 observers]

		63 inches	*67 inches*	*71 inches*	*75 inches*	*Mean (69 inches)*
14 yards	M	63.48	67.62	71.90	75.71	69.67
	S.D.	1.89	1.66	1.83	3.05	......
28 yards	M	60.44	66.76	70.84	75.88	68.84
	S.D.	5.12	5.16	5.71	5.16	......
56 yards	M	63.28	68.64	72.96	78.32	70.80
	S.D.	5.27	5.32	7.60	5.14	......
112 yards	M	64.92	71.56	75.20	78.60	72.57
	S.D.	7.21	6.84	5.84	6.18	......
224 yards	M	68.16	71.52	75.84	78.96	73.62
	S.D.	8.99	7.96	7.26	6.49	......
448 yards	M	67.68	70.76	74.08	77.24	72.44
	S.D.	10.07	9.93	8.93	8.55	......
784 yards	M	68.84	69.96	74.92	9.31	72.28
	S.D.	9.85	11.73	9.85	75.40	......

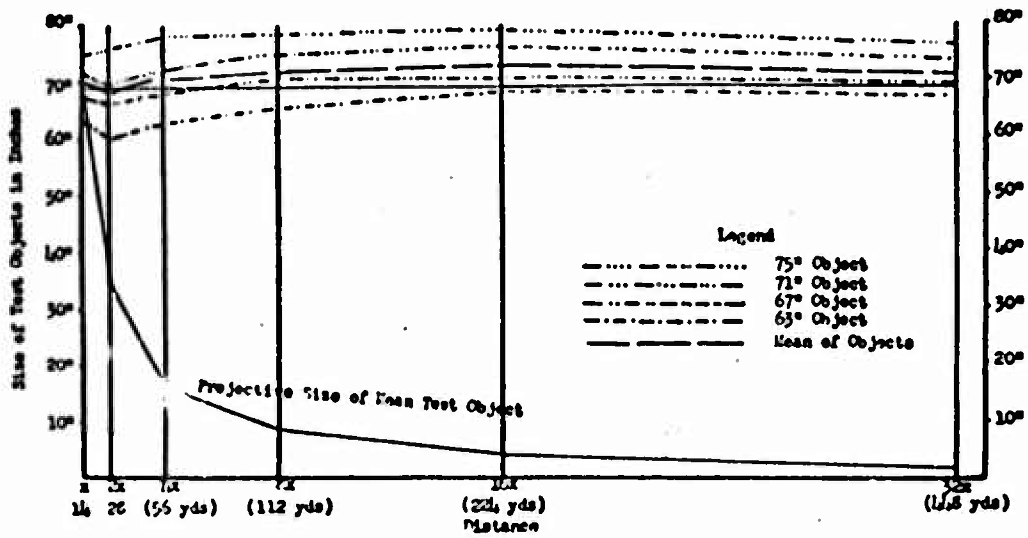


FIGURE 9.6.—Apparent Size of Distant Test Objects at Increasing Distances in Real Space. (The solid lines in the graph represent the *actual* size of the mean test object and the *projective* size of the mean test object. This refers to a theoretical object of 69 inches.)

figure 9.6, but reference to the table shows that they are essentially the same as at 448 yards. The variability of the perceived sizes rises somewhat with increasing distance, but the mean apparent size remains constant up to a point where the object is barely visible. In figure 9.6 the lower solid line represents the retinal or projective size of the mean test object and the upper

straight line represents its actual size, i. e., 69 inches. The perceived size follows the upper line and continues to do so although the retinal size diminishes and will at some further distance become imperceptible. The "amount" of size constancy is theoretically given by the extent to which the perceived size departs from the lower line and approaches the upper line. In this experiment, however, there is no question of the amount of size constancy. Constancy is simply the rule. The mean perceived size is, in fact, slightly greater than the actual size at all distances except 28 yards. There would be no value in computing ratios or indices of the amount of constancy in this experiment. A minority of observers have ratios somewhat under 100 and a larger number have ratios slightly over 100.

These results strongly suggest the conclusion that as long as one can see an object at all, it will be seen at approximately its true size, at least under conditions favorable for viewing distance. This conclusion is somewhat surprising on first examination and appears to contradict other studies. The observers in these other studies, however, were in all probability not making practical estimations of ~~size-at-a-distance,~~ with the realistic attitude which this implies. The conclusion is consistent with the theory of retinal gradients of texture and size of similar objects according to which retinal size and distance are reciprocal up to the vanishing of all sizes at the horizon. The elements and areas and objects of the terrain, as the flier sees it stretching away from him, do not appear to be smaller than they actually are as their distance increases. A distant landing strip, for example, will not be consistently underestimated in size as compared with a near-by landing strip. It will only be harder to estimate accurately.

The standard deviations of the mean perceived sizes in table 9.6 tend to increase with distance. This fact simply confirms the conclusion already reached that *accuracy* of size estimation, and presumably of distance estimation as well, declines with distance. The distinction to be noted is that although accuracy of size perception decreases, *constancy* of size perception does not.

TABLE 9.7.—*Means and sigmas of size estimations for each test object when presented at different distances in photographs*

[Based upon 5 judgments of each object at each distance obtained from each of 13 observers]

Distances		*63 inches*	*67 inches*	*71 inches*	*75 inches*	*Mean (69 inches)*
14 yards	M	63.4	68.4	72.1	76.1	69.9
	S.D.	1.22	1.98	1.86	2.10	......
28 yards	M	67.7	8.11	76.8	79.7	73.1
	S.D.	7.68	68.1	6.98	6.73	......
56 yards	M	68.8	7.54	77.3	80.0	74.9
	S.D.	7.73	73.6	7.56	7.57	......
112 yards	M	67.9	72.6	74.8	76.1	72.8
	S.D.	8.39	7.94	7.86	8.42	......
224 yards	M	66.3	69.3	71.4	75.3	70.6
	S.D.	7.88	8.85	8.54	8.46	......
448 yards	M	62.7	64.2	66.3	68.4	65.4
	S.D.	9.87	11.36	10.64	10.97	......

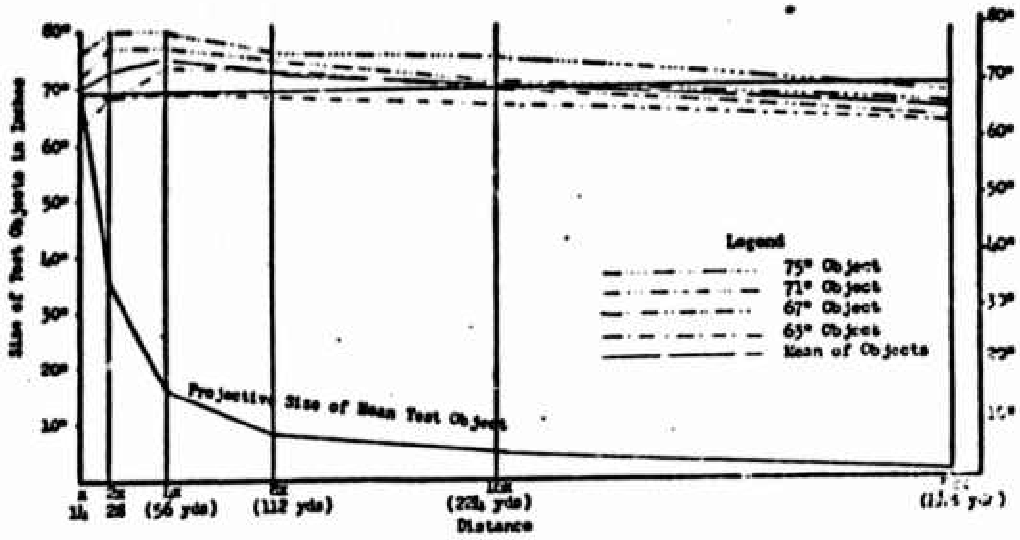


FIGURE 9.7.—Apparent Size of Distant Test Objects at Increasing Distances in Photographs (The solid lines in the graph represent the *actual* size of the mean test object and the *projective* size of the mean test object. This refers to a theoretical object of 69 inches.)

Table 9.7 and figure 9.7 show the corresponding results to those just presented for the experiment with photographic space. They confirm the general conclusions reached in every respect. It may be noted that the variability of the mean perceived sizes is greater with photographs than in the outdoor situation, as would be expected. The constancy or objectivity of perceived size, however, holds true even in a representation of three-dimensional space, and does so even at a "distance" of 448 yards where the test object is so minute on the photographic print that it can only just be made out. It is true that the plotted sizes show a slight downward trend, but they nevertheless do not fall far below complete size constancy. It is likely, since the same observers were used in both the outdoor experiment and the photographic experiment, that there was some practice-effect to improve the accuracy of the latter estimates. It could not have been great, since the observers were not aware of the results. The amount of effect may be judged from the data to be reported next, which were obtained with a different group of subjects.

Construction of the Distance Estimation Test Form A (CP212A)

Arrangement and Number of Items. The same photographs used in the experiment were also used for the construction of the Distance Estimation Test CP212A. A description of these photographs and the procedure followed in presenting them have been given in the description of the photographic experiment. In that procedure although the *size* of the test objects was varied in an irregular manner, the distance at which they apeared remained constant for a block of 25 trials. In presenting the photographs

as a test it was possible to vary the distance as well as the size of the test objects from one trial to the next in a random fashion.

It was necessary to reduce the number of trials (from 150 used in the photographic experiment) to a smaller number which would be practicable in a test. Forty trials could be presented (each picture being shown individually) in about 15 minutes. This number of trials was therefore used in the testing procedure. Evidence will be presented to show that 40 trials is sufficient to give the test good reliability.

A third modification of the procedure used in the photographic experiment was introduced to permit a standard I.B.M. answer sheet to be used for recording and scoring responses. Since there were in the original experiment oversized and undersized test objects, and since judgments of higher than 15 and shorter than 1 were permitted, a total of 17 responses was possible. In order to reduce the number of possible judgments to 15 so that a 15-place answer sheet could be used, the over and under judgments were not mentioned in the instructions, and the two pictures showing the over and undersized stakes at each distance were eliminated from the test. It was also not necessary to include the block of trials in which the test stakes appeared at the same distance as the standards (14 yards). This control had already demonstrated that size perception *not* at a distance is relatively accurate. The basic number of photographs utilized, therefore, was 20, four heights of test objects (63, 67, 71, and 75 inches) represented at each of five distances (28, 56, 112, 228, and 448 yards). By giving the 20 photographs in a random order and then repeating them in another random order, a test of 40 trials was constructed, the two halves of which were equivalent.

Instructions. Preceding the test, the following instructions were given:

This is a test of your ability to perceive distance. In the test you will see photographs of a large field stretching away from you for a long distance and finally merging into hilly ground near the horizon. Close to you, in the foreground of the photograph, is a set of 15 white stakes. At a greater distance from you, in the background, is a single white stake. Each trial of the test will consist of a photograph. All the photographs are the same except that the height and the distance of the far stake will be different from one trial to the next.

Look at the photograph on the stand.

The 15 stakes in the foreground may be thought of as a scale of heights (extending from 1–15). Your task is to estimate the height of the stake in the distance in terms of this scale. If you could walk out to the distant stake and measure it, you would always find it to be exactly the same height as one of the 15.

Since the scale is near and the stake to be judged is distant, the accuracy of your estimate will depend indirectly on how clearly you see the distance to the far stake. But do not try to estimate this distance by itself;

702313 47--15

concentrate instead on how high the far stake looks. If you can get a clear impression of that you will have taken account of the distance naturally and automatically.

In some trials you will find the exact point of the scale difficult to determine. Do not expect your estimate to be perfect or exact. Your score will take into consideration not only perfect judgments, but also the closeness with which you approximate the real size of the distant stake.

You will have three practice trials before the test begins.

Practice Trial 1. (Distance 28 feet, size of test stake 75 inches.) Look at the distant stake. Which of the numbered series of 15 stakes directly in front of you is the same height? You may notice that the stakes in the foreground have slightly different widths. Disregard the width of the stakes in making your judgment; the width of the distant stake may be different from that of the stake which matches it in height. There is no way by which you can compute or measure the size of the distant stake. Your first clear impression of it is likely to be the most accurate one.

Now record the height of the distant stake using the numbers from 1 to 15 opposite item 1 on your answer sheet. (Examiner waits until answer is marked.) The correct answer to trial 1 is 13. The distant stake is actually height 13 in the scale. Correct your answer if you have it wrong. (Examiner waits if necessary.)

Practice Trial 2. (Distance 112 feet, size of test stake 63 inches.) Note that the far stake in this photograph is at a different distance from you than was the stake in the first practice trial. Record your estimate opposite item 2 on your answer sheet. (Examiner waits.) The correct answer to practice trial 2 is number 10. Correct your answer if necessary. (Examiner waits if necessary.)

Practice Trial 3. (Distance 56 feet, size of test stake 71 inches.) Again notice that the distance from you to the far stake has changed. Record the height of the distant stake opposite number 3 on your answer sheet. (Examiner waits.) The correct answer to practice trial 3 is 12. Correct your answer if necessary. (Examiner waits if necessary.)

The test consists of 40 trials. Do not hesitate too much over your estimates but do them as accurately as you can. Now, we will begin the test.

Scoring Formula. The formula used to score the tests was $3R + 2W_1 + W_2$ in which R = number of correct matches, W_1 = one-step wrong on the scale, and W_2 = two or more steps wrong. The reason for a weighted formula is that responses are not either *right or wrong.* If wrong they are *wrong by a certain amount* in terms of a scale of standard heights. It is, therefore, desirable to take into account the amount of error as well as the number of errors. An error is a deviation of the judgment in either direction. The specific weights used in the scoring formula were determined by an examination of the distribution of the correct matches and of nearly correct matches on the scale of standards. The simplest ratio which fit this distribution was 3:2:1. Evidence will be presented in the results to show that the reliability of the test is higher when this formula is used instead of simply scoring the number of correct matches.

Administration of the Test. The test was administered individually to each of a group of 50 subjects. They were all AAF

"returnees" who had been overseas and returned to Redistribution Station No. 4 at Santa Ana Army Air Base for reassignment or discharge. They were tested after they had received routine physical examinations by the Flight Surgeon. All had normal vision and good health. The majority were aircrew personnel (pilots, bombardiers, navigators, and gunners) who needed accurate distance perception in order to perform their duties.

Each man was tested individually in a well-illuminated room. Special care was taken that the light would fall evenly upon the surface of the photograph so that distracting highlights and the surface qualities of the print would be minimized. The photographs were presented individually in a holder which was placed at eye level about one foot away from the seated subject, and each print was made so that the scene appeared as if through an oval window. This was the best that could be done under the circumstances to enhance the depth effect of the photographs. The test was administered similarly to what would be possible with a printed test booklet of photographs, under group-testing conditions. No time limit was imposed on each trial of the test, but the subjects were asked to pace themselves rapidly without being careless in making their judgments.

The total time required to take the test (including instructions) did not exceed 20 minutes. The average time required was 15 minutes and some subjects finished in 12 minutes.

Reliability. It will be recalled that the test consisted of two identical sets of 20 photographs each presented in random order. Hence, the reliability of the test may be determined by correlating the two equivalent halves of the test. Two reliability coefficients were calculated: one betwen the number correct, or unweighted scores, of each individual on the two halves of the test, the other between the weighted scores based upon the formula $3R + 2W_1 + W_2$. These correlations are in table 9.8. From these

TABLE 9.8.—*Reliability of distance estimation Test (CP212A)*

Date	*Place*	*Group*	*N*	M_a	M_b	SD_a	SD_b	*r*	r_c	*Type*	*Remarks*
6/45	SAAAB	Returnees	50	3.04	3.20	2.05	2.04	.40	.57	Split half	Rights
6/45	SAAAB	Returnees	50	24.24	25.08	9.80	8.08	.66	.79	Split half	$3R + 2W_1 + W_2$

half-test reliabilities, the self correlation of the whole test of 40 trials was estimated by means of the Spearman-Brown formula.

The corrected coefficient of .79 based upon the weighted formula shows that the distance-estimation test is moderately reliable. This is the fact, despite the impression frequently reported that judgments of size-at-a-distance cannot be made with a subjective feeling of certainty. This finding confirms the high consistency in magnitude of size judgments at different distances in the photographic experiment. The reliability of the test is considerably

higher when the weighted scoring formula is used than when the test is scored by taking the number of right matches.

Distribution. The means and sigmas of the distributions of the unweighted scores (number correct) and the weighted scores are presented in table 9.9. When the test is scored on the basis of

TABLE 9.9.—*Distribution constants*

Date	*Place*	*Group*	*N*	*Mean*	*SD*	*Remarks*
6/45	SAAAB	Returnee (Air crew).....	50	6.42	3.42	Scored Rights.
6/45	SAAAB	Returnee (Air crew).....	50	49.26	16.4	Scored $3R + 2W_1 + W_2$.

number right, the mean is only 6.42 and the standard deviation 3.42. The test is too difficult. A more satisfactory distribution of scores for test purposes is obtained when individuals are scored with the weighted formula. The mean is then 49.26 and the standard deviation is 16.4. It will be noted that individual differences in this form of the ability to judge distance are pronounced despite the fact that the subjects are all fliers who have had the opportunity to learn judgment of distance in aerial long-range space.

Accuracy of Size-Distance Perception in Form A of the Distance Estimation Test. The accuracy of size perception at increasing distances in the test follows the same trend that was found in the experiment to validate the photographic method. In the test, as in the experiment, a decrease in the accuracy of perception occurs at each increase in distance of the test objects from the observers. Evidence for this conclusion is demonstrated by the successive increase in the mean of the average errors from 6.85 inches at 28 yards to 12.90 inches at 448 yards (table 9.10).

TABLE 9.10.—*Accuracy of size-at-a-distance judgments in the distance estimation test, Form A, as represented by average errors in inches*

[Based upon two judgments of each object at each distance from each of 50 observers]

Distances	*63 inches*	*67 inches*	*71 inches*	*75 inches*	*Mean (69 inches)*
28 yards	6.79	7.18	5.92	7.61	6.85
56 yards	7.35	6.86	7.28	7.84	7.33
112 yards	7.29	6.64	7.18	9.32	7.61
224 yards	9.11	8.44	8.36	9.50	8.85
448 yards	10.95	12.48	13.69	14.60	12.90

However, the accuracy of perceived size-at-a-distance in the test is somewhat less than it was in the experiment on the photographic method. The relative accuracy of size perception at a distance in the test and the experiments may be obtained by comparing tables 9.1-A and B and table 9.10.

Constancy and Variability of Size-Distance Perception in the Distance Estimation Test. The means and standard deviations of the size estimates of each object at each of five distances are presented in table 9.11. These results show that the predominant tendency in the test is for subjects to judge the real size of the test objects rather than the image size on the photographs. There

TABLE 9.11.—*Means and standard deviations of perceived sizes of objects at various distances in the distance estimation test, Form A*

[Based upon two judgments of each object at each distance from each of 50 observers]

Distances		*63 inches*	*67 inches*	*71 inches*	*75 inches*	*Mean (69 inches)*
28 yards	M	59.60	69.76	71.40	69.76	67.63
	S.D.	8.31	8.68	7.71	8.77	......
56 yards	M	60.40	67.16	67.08	68.88	65.88
	S.D.	8.52	8.85	8.62	8.52	......
112 yards	M	60.88	66.64	66.96	67.12	65.40
	S.D.	8.56	7.97	8.34	8.81	......
224 yards	M	59.32	63.0	65.80	67.24	63.84
	S.D.	10.59	9.68	10.35	9.60	......
448 yards	M	58.36	58.60	60.76	61.16	59.72
	S.D.	13.50	12.74	13.94	11.94	......

is, however, a slight reduction in constancy with each increase in distance represented in the photographic test. At 28 yards, the mean size estimate for an average object of 69 inches is 67.63 inches, at 224 yards, 63.84 inches, and at 448 yards, 59.72 inches. This slight trend toward under-constancy, small at first, but increasing from 224 to 448 yards is in contrast to the over-estimation of size at great distances found in the results of the experiments given in an earlier section (*cf.* tables and figures 9.6 and 9.7). Table 9.11 also shows that the variability of judgments increases with each increase of distance. This increase is somewhat larger than that found in the earlier experiment.

These results verify the use of the test as a measure of size-distance perception. Perception of size decreases in accuracy with increasing distance in the test, as was the case in a natural outdoor situation. Constancy is maintained to a considerable degree at all distances in the test as it was outdoors and also in the photographic experiment. Variability of the size estimations increased also in the test. To the extent that the test results are comparable to those obtained in the experiment, it may be stated that the photographs presented in a random test order do indeed measure a tridimensional spatial ability, the perception of size-at-a distance The differences that exist between mean accuracy and constancy of judgments in the test and the experiment simply imply that the test situation was not as favorable to the judgment of size at a distance as was the experimental situation. This is to be expected also on statistical grounds, since only two judgments of each size at each distance were made by individuals taking the test, whereas five judgments were obtained in the experiment.

Construction of the Distance Estimation Test, Form B (CP212B)

A partial revision of the test just described, numbered CP212B, was planned but the new photographs required for it were not completed. It was considered useful to revise the original test for a number of reasons.

1. The level of difficulty of the test was probably higher than is desirable, resulting in a relatively small percentage of actually

correct judgments. This was taken care of by using a weighted formula. It would be advantageous, however, to construct an easier test which it might be possible to score simply by counting the number of correct answers, i. e., a test in which the steps of the scale are greater.

2. In Form A, the four test objects whose size had to be judged corresponded to the stakes numbered 10, 11, 12, and 13 in the standard scale of 15 sizes. They were chosen from one end of the scale in the expectation that most of the errors would be underestimations of size requiring a relatively larger number of stakes smaller than the test object. The results proved this expectation (based upon evidence from laboratory constancy experiments) to be false. In fact, the frequency of *overestimations* of size occurring in Form A makes it desirable to have an equal number of standards on both sides of the test objects. This symmetrical arrangement of the test objects in relation to the standard scale was adopted in constructing Form B. The test objects were also selected over a wide range of the scale.

3. Photographic reproduction of the stakes at 448 yards in Form A was close to the limit of their distinguishability. Careful photographic processing is necessary in order to produce adequate prints of the test object at this distance. Although observers agree on the relative sizes of the test objects at this distance, if the test were to be reproduced many times in booklet form, it would be desirable to modify the scale of distances, particularly by decreasing the farthest distance.

4. Since 15-place I.B.M. answer sheets lettered from A to O are most easily available, it would be advantageous to letter the standard stakes instead of numbering them.

Form B should be photographed in an outdoor environment similar to that employed in Form A. The scale of standards consists of 13 white stakes set into the ground in an arc. They are 18-22 inches apart, with a larger gap of 6 feet in the center of the series. The standards are arranged in order from the shortest to the tallest and are lettered consecutively from A through M. The shortest standard on the left is 25 inches high and the stakes increase in size by five-inch intervals (instead of 4-inch intervals) up to the largest stake which is 85 inches. These standards may be seen in the foreground of the photograph; their actual distance from the camera is 12½ yards.

Through the central gap between stakes lettered F and G is seen the test object in the distance. Four different heights of test objects, 35, 50, 65, and 75 inches in size are shown at six different distances, 25, 50, 100, 200, 300, and 400 yards, in different photographs. The basic number of photographs in Form B is therefore 24. By random arrangement of two sets of these

photographs, a test of 48 trials can be constructed. This will have an administration time of approximately 15 minutes. According to evidence obtained on Form A its reliability can be expected to be about .80. The directions are similar to those for Form A.

The following are the principal modifications of Form A incorporated in Form B:

1. Thirteen lettered stakes with progressive increases in size of five inches are employed as standards instead of fifteen numbered stakes which increase in size by four-inch increments. The larger increment is expected to lower the difficulty level of the test.

2. The test objects are 35, 50, 65, and 75 inches in height, as compared with 63, 67, 71, and 75 inches. They are not grouped together in a restricted part of the scale but are distributed throughout the scale. Since there are an equal number of standards larger and smaller than the test objects, overestimations can be measured as effectively as underestimations. The test objects should be photographed in horizontal as well as vertical positions so that either type of position may be used in the test.

3. The farthest distance should be reduced from 448 yards to 400 yards in Form B. An additional distance of 300 yards should be introduced. Either or both of these distances may be used in the test, depending upon the clarity of the most distant stake as reproduced for test purposes.

THE ABILITY TO JUDGE DISTANCE AND SPACE IN TERMS OF THE RETINAL MOTION CUE

The distance perception heretofore discussed has been that of a stationary observer. The flier, however, is in motion. It is primarily because he is in motion that distance perception is so important to him. The stimulus of retinal motion perspective as a basis for distance has been described only in part. Its nature and application to the perception of aerial space during flight may next be considered.

Types of Retinal Stimulation in Relation to Visual Motion Perception

The retina may be stimulated by motion, or more exactly, the retinal image may undergo motion, in two general ways. The first and simplest is relative displacement. In this case the image corresponding to an object is displaced in relation to the image corresponding to the rest of the world, i. e., the image of the background against which the object-image is located. This is the stimulus that exists when an object in the visual scene moves. The relative displacement is the same whether the eyes follow the moving object or not. The second may be called deformation. In

this less familiar type of stimulation, the image is distorted as a whole; it stretches or expands or contracts rather than merely being transposed. This sort of change obviously does not physically occur in real objects, or at least solid objects, but it happens all the time to our retinal images—especially to the image of the terrain or background. The fact to be especially noted is that the retinal image over the whole retina, the image representing the entire visual field, may undergo this kind of motion, i. e., may flow at different rates in different parts.

Having distinguished between retinal displacement and retinal deformation, the kinds of visually perceived motion which correspond to them can be stated. The general rule may be formulated that whenever the *observer himself* moves, the retinal image corresponding to the whole visual field undergoes deformation. The converse is also true. When the observer's body is motionless, there is *no* deformation of the retinal image as a whole.

When *objects* move, the corresponding object-images within the retinal image of the field undergo relative displacement (and may also undergo deformation if the objects move toward or away from us) but the retinal background-image of the whole field does not undergo deformation. This rule holds even though the eyes may move from one fixation to another or may fixate a moving object, *so long as the head does not move*, i. e., so long as the *position* of the eyes in space does not change.

When *both* the observer and objects move in a three-dimensional space, there occurs both deformation of the retinal background-image, and displacement (possibly with deformation) of the retinal object-images. Both the observer's own movement and the movement of objects are perceived simultaneously, under normal circumstances, without any interference between the two kinds of perception.

The importance of these distinctions lies in the fact that deformation of the retinal background-image yields not only the perception of subjective motion but provides a powerful stimulus for space perception. Different aspects of this deformation are specific not only to the direction of one's motion and to the velocity of one's motion, but also to its angle of inclination of the surface at which one is looking, to the distance of all points of the surface, and in fact to the distance of all stationary objects in one's field of vision. Considered as a stimulus-correlate for distance, the flowing deformation of the retinal image is identical with what we have called retinal motion perspective.

In order to avoid complicating the discussion unnecessarily one assumption should be made at this point. For the sake of simplicity and for the time being, it will be assumed that the eyes of the moving observer are always fixated at an infinitely distant point

such as the horizon, and therefore motionless in his head. In other words we will disregard movements of the eyes, particularly pursuit movements. The retinal displacement produced when the eyes of a moving observer themselves rotate can be considered independently and therefore the problem can be deferred.

Retinal Motion Perspective. When the observer moves, the retinal image flows at different rates on his retina in an exactly inverse relation to the distances of the corresponding points or objects in the world. The more distant the point, the less its velocity. The line of the horizon, for example, does not move on the retina. Objects at distances which are for practical purposes infinite will not move either. Consider the appearance of the visual world on a clear night, when the observer is flying or driving a car through it. The stars, the moon, and the horizon do not undergo any deformation whatever; their distance is so great that the observer does not change position in relation to them or, in other words, the parallax is zero. But the terrain and the objects on it flow across the visual field, slowly at distant points and more and more rapidly at points nearer the observer. The differences in the rate of flow are dependent on differences in distance, since the nearer points change their direction from the observer, and are therefore retinally displaced, faster than the further points. There exists a certain degree of parallax which is inversely proportional to the distance. Considering this fact as a matter of different velocities stimulating the retina at the same moment, the phenomenon is what we have termed motion perspective.

In addition to the fact that the velocity of this retinal flow approaches zero for very distant objects and vanishes, therefore, at the horizon, it also approaches zero and vanishes at two specific points in the visual world—the point toward which the observer is moving and its opposite, the point he is moving away from. For example, it is a familiar fact that an object in the direct line of an observer's progress yields retinal motion only to the extent that it expands; the exact point on the object at which locomotion is aimed does not move at all. Optically speaking, therefore, the world expands radially outward as the observer moves into it, and, assuming he looks backward, contracts radially inward as he moves away from it. The expansion may be noticed when one drives a car on a straight road, particularly at night, and looks at a distant point ahead. The corresponding contraction is most easily observed from the rear end of a train.

Figure 9.8 shows diagramatically the general character of retinal motion perspective as it exists when the observer is moving and looking straight ahead over a level terrain. The arrows are vectors and the length of each arrow represents the retinal velocity

703313—47—16

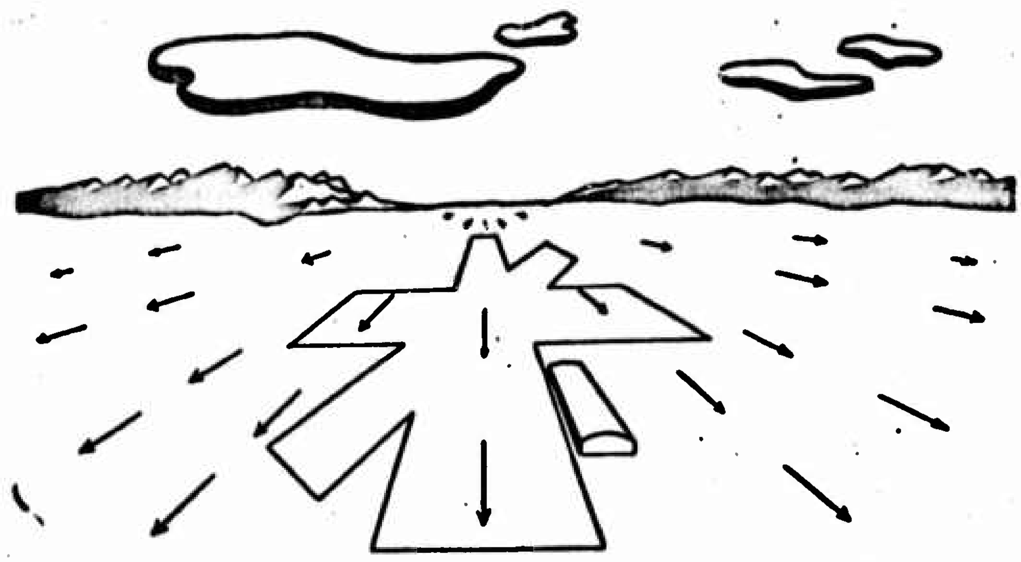


FIGURE 9.8.—Retinal Motion Perspective Looking Ahead.

in the field of view at that particular point in the field. These velocities exist, of course, as *continuous gradients* rather than as separate *displacements* on the retina. The retinal image is deformed as a whole. A gradient of velocity is a rate of change of velocity, not a velocity as such. The distinction is important, since the hypothesis will be advanced that the *gradients* in the retinal field and the direction of these gradients rather than the velocities themselves are the effective stimuli for the perception of distance, space, and locomotion.

In figure 9.8 it should be noted that there is a gradient of velocities varying from zero at the horizon to a maximum at the near region of the ground. From the observer to the horizon in any direction, this gradient of *velocity* is repeated. The *direction* of all retinal velocities is radially outward from the point toward which one is moving, i. e., from the "center of expansion." In figure 9.8 the center of expansion is on the horizon because the observer is assumed to be moving parallel to the terrain.

Figure 9.9 shows the motion perspective when the eyes are looking not ahead but to the right, as for example in looking out the side window of a plane or train. All velocities vanish at the horizon. They reach their maximum at the point directly under the observer, which point of course cannot be represented in the diagram. The vanishing of all velocities at the horizon is analogous in this diagram to the vanishing of all sizes on the horizon, and in fact the gradient of decreasing velocity upward on the flat visual field is similar to the corresponding gradient of retinal size. In this diagram, at least, retinal motion perspective is similar to size perspective and to linear perspective. Unlike size, however, motion has the characteristic of direction.

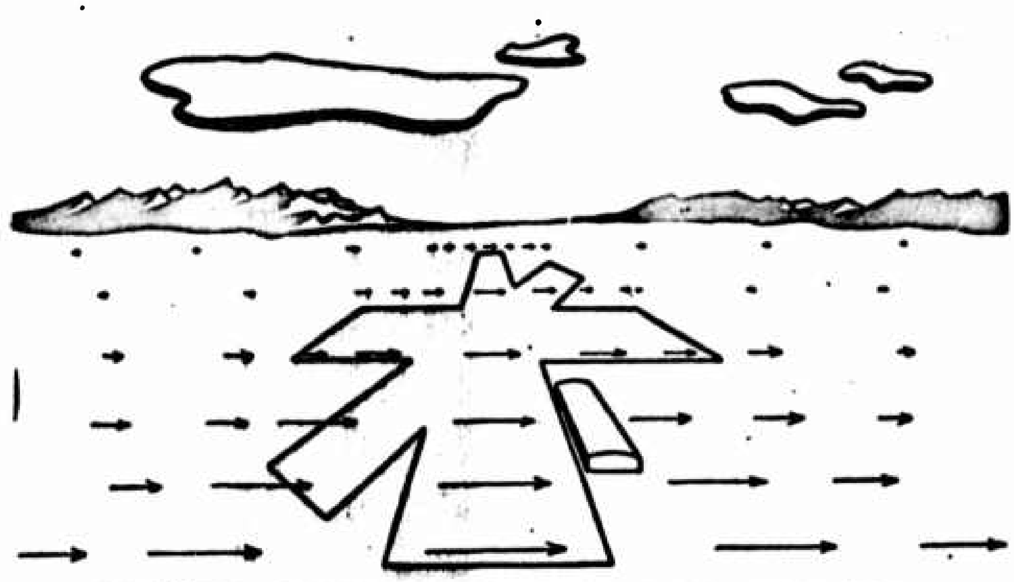


FIGURE 9.9.—Retinal Motion Perspective Looking to the Right.

If one were to *reverse the direction* of the arrows in figure 9.9 or in other words make the field move in the opposite way, the diagram would represent the view of an observer who is looking straight to the *left* of his line of motion instead of to the right. Similarly, reversing the direction of the arrows in figure 9.8, which would have the effect of making the retinal field contract instead of expand, would represent the view of an observer looking in the opposite direction to his line of motion, that is, looking backward. There are, therefore, two opposite poles in optical space at which retinal motion is zero, the center of expansion forward and the center of contraction backward on the line of the observer's locomotion. These poles exist independently of retinal motion perspective. They are cues for the perception not of distance but of the direction of locomotion.

The third diagram, figure 9.10, represents the motion perspective when the observer moves forward but looks vertically downward, as for example in looking through the bomb bay doors of an airplane. The velocity of the terrain going by is greatest at its nearest point—the one vertically below the observer. The velocities decrease radially and symmetrically from this maximum point and, merging with the family of gradients already described, finally diminish to zero at the encircling rim of the horizon. By visualizing all of these views as if they were continuous and connected, a fairly complete picture may be gained of the retinal gradients of motion perspective and their interrelations.

The answer to the question of how the retina is stimulated if the observer looks vertically upward into a clear sky during locomotion is probably obvious—there is no motion stimulus at all. If nothing is seen but sky there is no visual sense of one's own movement. Likewise there is no visual experience of a world of

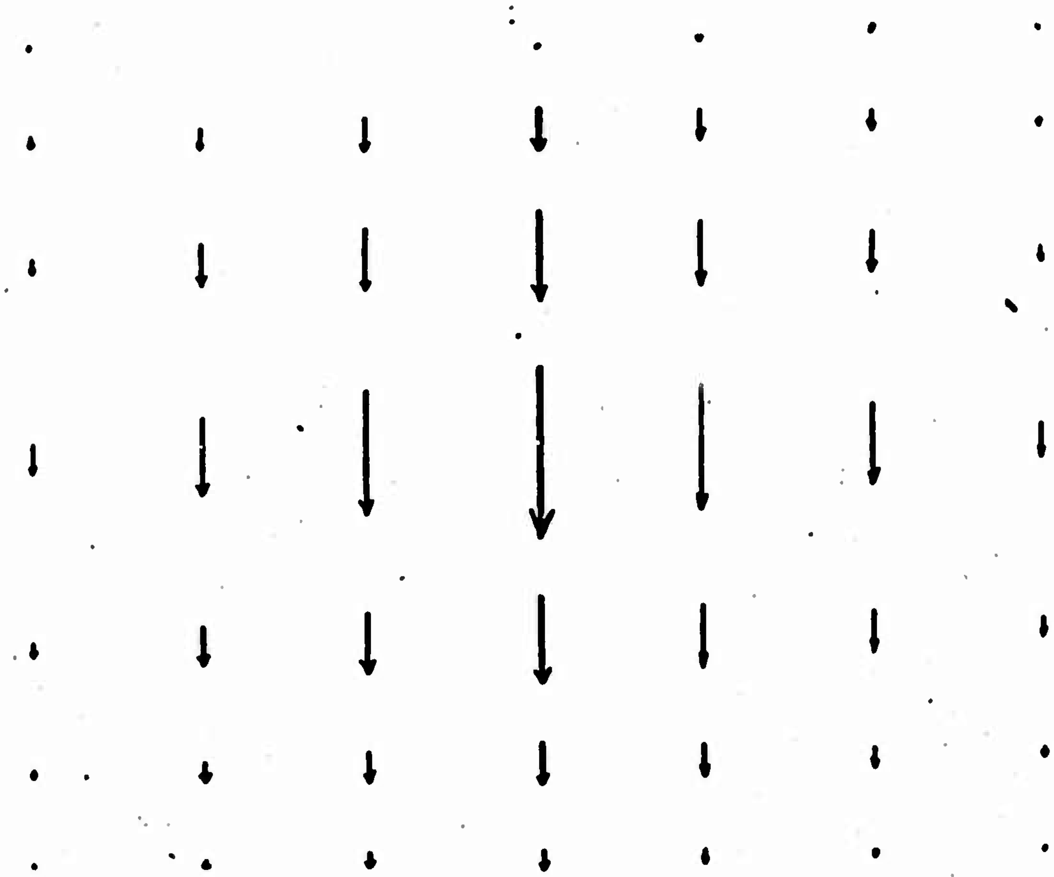


FIGURE 9.10.—Retinal Motion Perspective Looking Down.

objects in three dimensions. Motion perception presupposes the stimulus of texture—the existence of objects or surfaces producing a differentiated retinal image. Only when there is such an image can retinal displacement or deformation occur. If the retinal stimulus is perfectly homogeneous or uniform at all points, as it is when we see nothing but clear sky, no motion can be perceived. If there are clouds in the sky, the surfaces formed by them will show retinal motion perspective and if the clouds are near enough to the observer, as occasionally happens in flying, the motion perspective will be noticeable. Figure 9.11 gives the appearance of the world ahead when flying under a solid overcast or "ceiling." If the cloud is sufficiently solid to be a visible surface, the motion perspective will be precisely analogous to that of a physical ceiling.

The Effect of Eye Movements on Gradients of Velocity. It is now possible to consider the relation of eye movements to retinal motion perspective. When the eyes follow a moving object or rotate from one fixation point to another the retinal background image obviously shifts across the retina itself. This shift is a displacement of the image of the visual field, not a deformation of it. Suppose that the eyes are fixated not on the motionless horizon as we have heretofore assumed but on a nearby point on

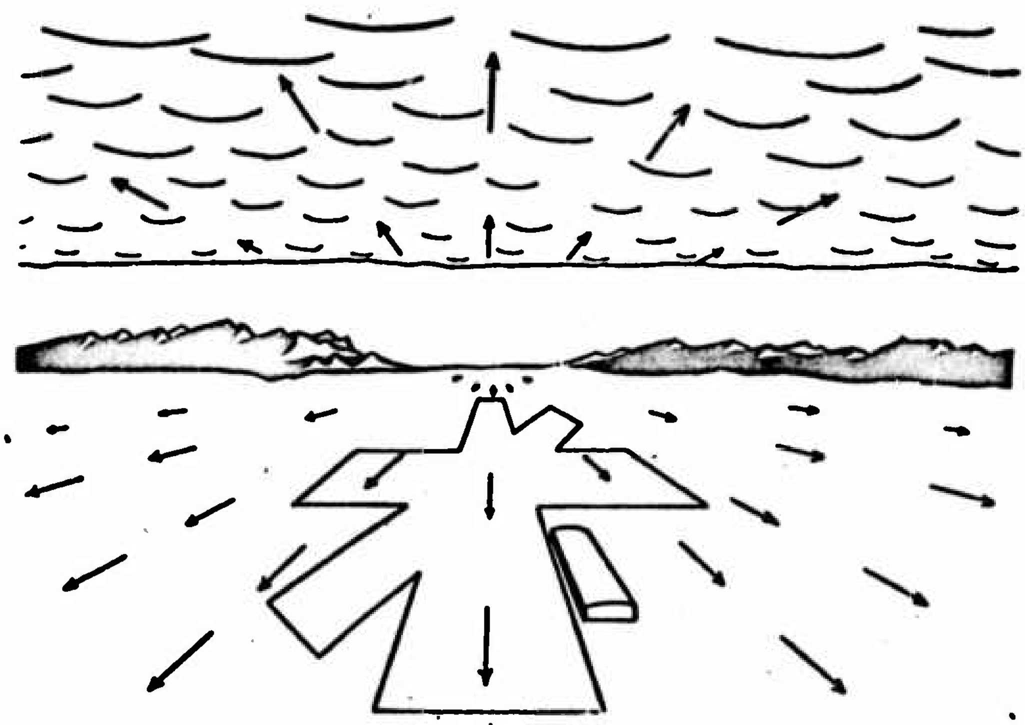


FIGURE 9.11.—Motion Perspective with a Ceiling.

the ground. During such a fixation, the retinal velocities in the field are quite different from those represented in the diagrams and one might suppose that retinal motion perspective is thereby destroyed. The retinal velocity at the point fixated is now zero; the velocity of that point has been just canceled by an opposite velocity induced by the rotation of the eye. The velocity at all other retinal points is affected by this compensatory velocity—it is in fact added to each f them. The horizon, for example, moves on the retina at the rate which the now fixated point formerly possessed. The fact is, nevertheless, that the observer sees essentially the same thing as if he looked at the horizon—a three-dimensional world in which he himself moves.

The explanation of this fact is that when a constant opposite velocity is added to a continuous series of different velocities, the relation between them remains unchanged. The adding of velocities is algebraic in the sense that a positive velocity in one direction may be canceled by a negative velocity in the other. The *gradient* of retinal velocities with respect to their direction is therefore unaffected by pursuit movement of the eyes. Since we know that the observer has the same experience of visual locomotion whether he fixates the horizon or watches the ground going by, we must suppose that the effective stimulus for such perception is the gradient of velocities in the retinal field—the direction and rate of change of velocities along a retinal axis—rather than the velocities themselves. The perceptual mechanism involved is

similar to that already discussed for the gradient of retinal disparity.

The Perception Resulting from Retinal Motion Gradients. The visual movement which has been described in the previous pages is the movement which occurs on the surface of the observer's retina—it is not of course the same thing as the movement which he sees or experiences. The aim has been to define the physiological and optical stimuli for his perceptions of space and locomotion. It is on the basis of these stimuli that he flies an airplane or otherwise gets around in an extended world. The observer does not "see" his retinal images, although his retinal images are of course the conditions of his seeing. Images are two-dimensional, whereas his perceptions are three-dimensional. The visual field which undergoes deformation, for example the expansion shown in figure 9.11, is the retinal field; the visual world resulting does not expand correspondingly but is seen as perfectly stable. Why does the world not expand when the stimulus image expands?

The answer must lie in the fact that the world possesses a third dimension and that the observer sees himself moving in this dimension. In all probability the expansion of the retinal image is an effective stimulus for perceived locomotion and for distance perception and by virtue of this fact it is *not* a stimulus for a perceived expansion of the visual world. This explanation suggests that if we could see the terrain of figure 9.11 as a flat image, we would also see it as one which instead of extending into the third dimension expands in two dimensions. If it can be assumed that retinal motion perspective during locomotion is normally a stimulus for distance, then it is only reasonable that it should not be a stimulus for deformation.

The customary terms to describe the fact that objects do not undergo expansion or contraction when their position is changed relative to the observer are "shape constancy" and size constancy." These phenomena are, however, probably only special cases of the general constancy of a continuous three-dimensional terrain.

Motion Perspective when the Observer's Movement is not Parallel to the Terrain; Application to the Problem of Landing an Airplane. Up until the present we have been dealing with the situation where the observer moves on the ground or parallel to the ground. This is the case during most of the ordinary small movements of human activity, and during walking, driving, and straight and level flying. However, the ability to judge distance and space is most critical not in these situations but in the situation where the observer is moving *toward* the ground as he is when landing an airplane. We are now in a position to describe the retinal motion stimuli of the pilot during a landing and the

way in which they are indicative of the direction, velocity, and angle of his glide, and of his distance or altitude from points on the ground.

The fundamental difference between the retinal gradients of motion during level flight and during a glide is that the optical center of expansion, the point toward which the observer is moving, is no longer on the horizon, but is on the ground[11]. Retinal motion therefore vanishes both on the rim of the horizon and at another point in the field—the center from which the motion radiates. Consider the case which is theoretically simplest—that of vertical descent with the observer looking straight down along the line of locomotion. The retinal velocity will be zero at the center of expansion, will increase to a maximum at points halfway between the center and the horizon, at 45° from the line of locomotion, and will then decrease and finally vanish at the horizon itself, at 90° from the line of locomotion. This pattern of the gradients of expansion will be symmetrical and will be quite different in character from the gradients existing during level flight.

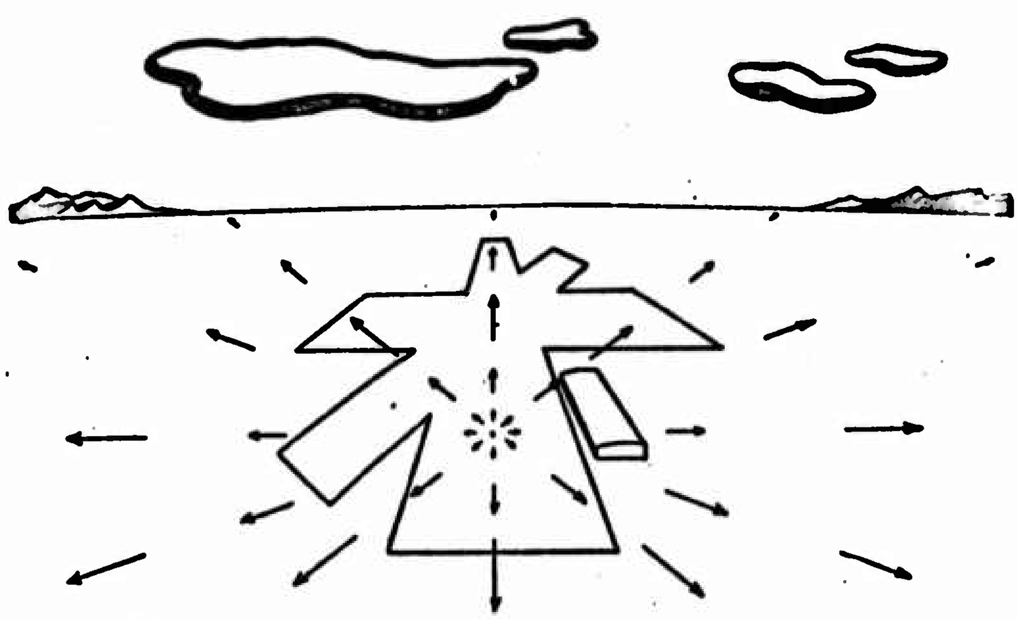


FIGURE 9.12. —Retinal Motion Gradients During a Landing Glide.

The gradients during a glide aimed at a point *between* the one vertically below and the horizon are a combination of the two situations described. They are diagramed in figure 9.12. The center of expansion is the point at which the glide is now aimed; if the glide is made steeper, it moves downward and if the glide is made shallower it moves upward toward the horizon. The center is, in other

[11]G. C. Grindley has described this and some other aspects of the phenomenon in a brief report to the British Flying Personnel Research Committee. Whether a discussion of the phenomenon has been published could not be determined at the time of writing this report.

words, an indicator of the direction of one's flight—of where one is going at the present moment. If the glide should be continued unchanged this zero point of the retinal gradients is the point where the wheels will touch ground. It therefore enables the pilot to see whether he is overshooting or undershooting the field. A plane, it should be noted, cannot be "aimed" by lining it up with a distant point like a rifle or even like an automobile; its nose never points exactly in the direction in which it is going but somewhat upward from this direction; in a cross-wind furthermore, the nose points to one side of the true direction of flight. The cue therefore by which flight can be guided or aimed has to be external to the plane itself and must be looked for in the visual scene ahead. Langewiesche[11] has described what is in effect the center of expansion cue in the terminology of the flier. He explains a method for estimating the landing point on the runway even before one turns into the approach by observing the center of expansion during the base leg, noting its angular distance below the horizon, and then projecting it in imagination on the runway, allowing for wind. He makes special note of the fact that, in a normal glide, what we have here called the center of expansion is always a specific angular distance below the horizon. This angle is, in fact, the characteristic gliding angle of the plane being flown. In the absence of power, that is in the event of forced landing, the point-of-present-aim cannot be made to rise above this angular distance; in other words the glide cannot be "stretched." The ability to judge this visual angle enables the flier to see at all times how far on the terrain he could glide with a dead engine.

The optical center of expansion is therefore an exact indicator of the direction of a glide and a means of seeing whether the present line of flight will or will not take the plane to the point the pilot wants to get to—a point, let us say, just above the near end of the runway. It is also an indicator of the angle of glide or the rate of descent. The angle of glide is in fact the visible angular distance between the horizon and the center of expansion. With an unfamiliar plane, or with unfamiliar wind conditions which may steepen or lengthen the glide, it becomes very important for a pilot to be able to see this angle of descent in the three-dimensional space in front of him rather than having to guess it or remember it from past experience.

In the landing situation of figure 9.12, the distance of the pilot from the center of expansion and his altitude from the ground directly below him need to be judged with a considerable degree of accuracy. What are the criteria or indicators of this distance and altitude? The coarseness of texture and the size of objects on the

[11] Langewiesche, W., *Stick and Rudder*, New York: McGraw-Hill, 1944, Ch. 16.

ground directly ahead are obvious correlates of these distances[11], as is increasing binocular depth or relief. A high tower or building on the airfield will provide a cue to the pilot's altitude; when the top of the tower just cuts the horizon, its height will be the same as his altitude. But the expansion gradients themselves vary concomitantly with distance from the ground and provide a basis for judging it. During a descent at uniform speed, the rate of the retinal expansion at its maximal points increases as the distance from the ground decreases. Under a given set of conditions this overall rate of expansion at any moment indicates the altitude. The fact that the ground directly below the plane goes by faster as the altitude decreases has been frequently described and discussed as a cue for landing. But this fact is only one aspect of the more general phenomenon of an increase in the rate of expansion over the entire field. The debated question as to how student pilots should use their eyes during landings ought to be considered in the light of the general phenomenon of expansion at *all* visible points.

The pilot's ability to estimate his ground speed both during landings and during straight and level flight, for what it is worth, is dependent wholly on the perception of retinal motion. We are referring now to ground speed rather than air speed, which latter is given by the air-speed meter, auditory cues, and others. At any given altitude, the over-all rate of the retinal pattern of gradients is directly proportional to his ground speed. Since the visual cue for speed and one of the cues for altitude are, if this analysis is correct, the same stimulus, it follows that visual judgments of speed and altitude are interrelated. The conclusion seems to be supported by the observation that speed appears to decrease as a plane gains altitude and increase as it loses altitude.

In summary, the retinal pattern of motion stimulation has aspects which are specific to a number of corresponding aspects of aerial space—ones which are of great significance to the job of flying. In addition to the dimension of distance on a continuous terrain they include the direction of flight and its angle to the ground, the altitude or distance above the ground and the velocity of flight.

The analysis of the retinal motion pattern that has been presented is admittedly theoretical. Some empirical verification is furnished by the motion-picture test derived from it which is next to be described, but that is insufficient. The theory needs to be formulated more exactly by the use of methods similar to those of projective geometry. Only a beginning has been made in such a task. Such functions as have been derived are omitted here in view

[11]A photographic test of ability to estimate altitude from the texture of the ground, at one time called Estimation of Altitude CP218A, was constructed at Psychological Research Unit No. 3 with some original assistance from the Psychological Test Film Unit. It was apparently never completed.

of their incompleteness. The generalizations that have been made have been checked, however, by geometrical methods or by empirical study of motion-picture images on a screen. A motion-picture image is, of course, analogous to a human retina insofar as two-dimensional gradients are concerned. The generalizations, therefore, are believed to be sound.

A MOTION PICTURE TEST FOR ACCURACY OF JUDGMENT DURING LANDING

Method of Constructing the Test

Preliminary Photography. In the constructing of this test, reported in Chapter 5, a considerable amount of preliminary research was carried out, over a period of nearly two years, on the problem of representing by motion pictures the appearance of the ground during an approach to a landing. A number of shots were made of a runway, and of other types of terrain, with the camera mounted in the nose of an airplane during a glide, the camera being pointed as nearly as possible along the axis of the glide. These scenes were studied and the existence of the "expansion-pattern" on the motion picture screen was empirically verified. It was also demonstrated that such scenes give to the onlooker a realistic and uniformly compelling experience of being himself moved downward toward the ground in a slanting path, however much he may still be aware of sitting in a room and looking at a screen. This fact is not surprising since in most respects the camera records and the projector represent exactly the real visual situation, i. e., the same visual stimuli that the onlooker would have had if he himself had been where the camera was. But in at least two respects the representation is not exact. The cues for "depth perception" yielded by binocular (as contrasted with monocular) vision are absent, and, what is more serious, the visual field of the camera and projector is very different from that of the eyes. Whereas the eyes can "see" over an angle in the neighborhood of 170° horizontally and 130° vertically, the widest angle of view that can be registered by a motion camera is 49° horizontally and 37° vertically. This angular field of view of the *visual world* should not be confused with the visual angle of intercept of the *screen image* for an individual seated in the classroom. Evidence has been accumulated by the Psychological Test Film Unit to indicate that what one sees on the screen is, within limits, independent of the angular size of the screen image. No matter where one sits one sees what the camera saw. The camera, however, can see only a restricted field.

This restricted field of view in the motion picture situation, which fails to register the appearance of the ground at the outer periphery of the visual field in the real situation, might be expected

to detract very strongly from its reality. But actually it appears not to do so. The shots were reported by experienced pilots to represent the landing situation very adequately, and to induce the experience of actually "being there," at a visibly changing altitude, at a specific speed and at a certain angle of glide. The performance of the pilot in the camera plane could actually be criticized while viewing the scene. The only complaints about unreality of the scene arose from the fact that the slight yawing and pitching of the plane is much more noticeable in the motion picture situation than it is in the equivalent real situation. The irregular motion of the motion picture frame in relation to the ground is more obvious than would be the equivalent motion of the plane's windshield.

Originally, the intention was to construct a test of altitude judgment based upon the fact that altitude may be estimated (if the ground speed is constant) by the expansion of the ground. However, the scenes obtained did not satisfy the specifications for a test. What was needed was a set of shots in which the altitude, angle of descent, and velocity of motion were controlled. It was, therefore, necessary to employ model photography of a miniature set. The judgment to be made was also changed and the examinee was required to judge his point of aim on the ground, given an already established glide. As has been described, this judgment is dependent upon the accurate perception of the expansion gradients of the ground. The aiming point, in a landing, is the point of no motion in a terrain which is everywhere else expanding outward.

Final Model Photography. In motion-picture model photography, the result is indistinguishable from that of the full-size reality if all action is maintained to scale. This fact is consistent with our theory of the perception of *gradients* of size for space-perception instead of *absolute* sizes. After some computation, a 60 x 90-foot airfield was built at the studios of the AAF Motion Picture Unit on a scale of 1 to 48. On the first third of the runway were painted five lettered spots, one foot apart, lettered A, B, C, D, and E. A camera track was constructed, 30 feet in length, at an angle of 30° to the ground. This track could be set in five positions such that the path of the camera could be aimed at each of the five spots. In these original scenes, the *axis* of the camera was also lined up with the appropriate spot so that the spot would appear in the center of the picture frame. The camera was then moved at a constant speed of 1.2 feet per second, guided by the track, so as to represent a quarter-mile glide (1440 feet) down toward the runway. This glide was steeper and slower than that ordinarily obtained with any military airplane, but for purposes of the test this feature was believed desirable. The use of the

track also eliminated most of the unstable motion of the picture frame in relation to the ground which was so apparent in the real situation. Each of the five glides was photographed, at three different shutter speeds of the camera (36, 27, and 18 frames per sec.) so as to yield apparent speeds in the neighborhood of 26, 35, and 52 miles per hour. The lens used had a focal length of 35 mm. which gives a field of view 35° wide and 26° high. Fifteen basic scenes, 20 to 40 seconds in duration, were available for preliminary try-out and for construction of items for the test. The basic scene is represented by the three successive photographs of the landing field shown in figure 5.3 in chapter 5.

Administration in 35-mm. Form

Overcoming Artificial Cues. The fifteen basic scenes on 35-mm. film were arranged in random order with reference to the speed of glide and to the aiming point (A, B, C, D, or E). They were then administered to several groups of aviation cadets to obtain information for the specifications of the final 16-mm. form of the test.

One serious defect in the original scenes had to be overcome, namely the fact that the aiming-point appeared in the center of the frame on the screen. It was, therefore, possible to select the correct spot by estimating the center point of the picture rather than by finding the motionless point. Such a cue is artifactual in the sense that it is not similar to those used by pilots in making landings. It was, therefore, eliminated by offsetting the picture upward or downward in relation to the frame of the picture. The fifteen scenes were off-centered in a random order, seven being off-centered upward and eight downward in relation to the frame. Thus, the aiming spot was never in the center of the screen, but always displaced in an upward or downward direction.

An experiment demonstrated the tendency of individuals to select the spot in the center of the screen, and also that when this cue was eliminated, the difficulty of discriminating the aiming point, on the basis of the expansion cue alone, rose to a level appropriate for a final form of the test. Two forms of the test were administered to small groups of preflight cadets. In the first form, the aiming point was always in the center of the screen; in the second form, the correct spot was off-centered either upward or downward, by means of a specially-constructed framing device. A comparison of the percent of correct judgments on each of the five spots in these two forms is given in the following table:

Correct spot	*Centered (N = 45)*	*Off-centered (N = 66)*
	Percent	*Percent*
A	73	46
B	62	38
C	67	40
D	49	37
E	74	34
Average:	65	39

Over 80% of the first group were aware that they had selected the center of the screen as the aiming point. On the other hand, in the group shown the correct spot off-centered, fewer than 5% reported using the "centering" cue, which was in this case misleading.

The Development of the Script. Preceding the administration of the randomized 15 basic scenes, brief instructions were given to the students. These instructions simply consisted of describing the test, and pointing out the task required. Brief mention was made of the expansion pattern and the students were asked to find the aiming spot by finding the spot which showed no motion. However, it became obvious from the low scores obtained from the first groups that detailed instructions and practice were needed. After considerable revision a script was finally developed which proved adequate in this respect. It will be shown that the script is, in effect, a training period for learning to recognize and use the point of no motion in judging the aiming point. The script is given, in part.

Script for Landing Judgment Test, Form E

Main Title:	**Aviation Cadet Testing Program (Air Corps Symbol)**
Fade to:	**Landing Judgment Test—Form E (CP505E)** **Psychological Test Film Unit, Santa Ana Army Air Base, July 1944**

ACTION:	**VOICE:**
A rolling title appears on the screen over a background of a still scene of the airfield as it appears at the beginning of a glide.	This is a test of your ability to judge the direction of your approach glide toward a runway. You can tell this direction if you can determine the spot on the runway at which your glide is aimed. In this test you will see a runway from the cockpit of your plane. On it are located 5 spots, separated by equal distances, and lettered A, B, C, D, and E. One of these 5 spots will always be the aiming point of your glide.
Still shot in background changes into an approach glide toward runway. Rolling title continues until voice ends, then disappears, leaving the end of the glide on the screen.	Consider yourself at the controls of the plane as it approaches the runway. You are gliding in a perfectly straight course toward one of the 5 spots. (Pause) You are to decide which of these spots is the aiming point of your glide. Disregard the fact that you would ordinarily level off; simply pick out the point at which you are now aiming. (Pause)
IBM answer sheet shown on screen. Hand with pencil demonstrates method of filling in answers.	The test will consist of a series of trials in which your glides will have different aiming points. After each

ACTION:	VOICE:
	trial you will be given time to record your judgment on the answer sheet which you have before you. The letters A, B, C, D, and E on your answer sheet correspond to the 5 lettered spots on the runway. Fill in the space under the letter corresponding to the aiming point of your glide.
Pencil points to numbers.	Be sure that the number of the trial in the test corresponds to the number of the trial on your answer sheet.
Rolling title over still scene of runway as it appears at the beginning of the glide.	You will now be shown 5 practice trials in which you will have an opportunity to learn how to make these judgments. After each of the practice trials the correct answer will be given to you. Begin at number one on your answer sheet and fill in the appropriate space for each trial.
Title, "Trial 1," for 2 secs.	Trial One.
Slow glide begins toward spot C and continues until end of paragraph.	To make an accurate judgment, you must observe one very important *cue*, which can always be used in landing a plane. This cue consists of the fact that as you glide downward the ground appears to enlarge. There is, however, *one central point* around which the enlargement occurs. This is the *aiming point* of your glide. All other points on the ground appear to radiate away from it. In this particular approach, spot C is the center of enlargement, and is therefore the correct answer.
Title, "Trial 2," for 2 secs.	Trial Two.
Slow glide toward spot B synchronized with voice.	Watch the enlargement of the ground as you approach the runway. If you look closely you will notice that the whole of the field of view out of the edges of the screen is expanding. At the edges this expansion is most noticeable. It becomes slower and less noticeable toward the middle of the field, until at the aiming point itself, the ground seems to move directly toward you, but there is no expansion in any other direction. Direct your attention to the appearance of the spots on the runway and find the one which is the center of the expansion. (Pause) B is the correct answer for Trial 2.

ACTION:	VOICE:
Title, "Trial 3," for 2 secs.	Trial Three.
Slow glide toward E, synchronized with voice.	At the top of your glide the rate of expansion is relatively slow all over the field, but as you come closer to the runway note that the rate of enlargement becomes generally more rapid. It therefore becomes progressively easier to find the aiming point the closer you come to the runway. Notice that spots A, B, and C are expanding. A slow outward expansion of D is also perceptible. So your glide must be aimed at E, which is the center of enlargement. Notice that the center of enlargement is *not* at the midpoint of the screen; be sure you are not misled in making your judgment on that basis.
Title, "Trial 4," for 2 secs.	Trial Four.
Slow glide toward spot D, synchronized with voice.	At this point in your glide you are too far away to isolate the aiming point. However, you can eliminate from consideration those spots that clearly appear to be moving outward. They are the ones that are farthest from the center of the enlargement. Notice that A and D are clearly moving away from the center. Shift your attention to C, and finally to E, observing as you come closer that both C and E are moving slowly away from the center of enlargement. Spot D is therefore the aiming point and the correct answer. Notice that D is not the midpoint of the screen.
Title, "Trial 5," for 2 secs.	Trial Five.
Slow glide toward spot A, synchronized with voice.	The outward expansion of spots D and E is obvious. (Pause) C also moves away from the center, but at a slower rate. (Pause) As you get closer to the runway a very slow expansion of spot B is seen, while A remains motionless. Everything on the screen is radiating outward from A. (Pause) Spot A is therefore the one toward which you are aiming. Notice that A is not the midpoint of the screen.
Instructions appear as a rolling title over a still scene of the runway.	You will now be given 5 *more* practice trials. At the beginning of each scene try to narrow down your judgments to those spots that show the

ACTION:	VOICE:
	least enlargement; as you approach the end of the scene eliminate other spots, and at the last moment, choose the spot which shows no expansion. Do not make the mistake of fixing your eyes on one spot and keeping them there. Keep your eyes moving and look for the spot from which all other points on the screen radiate outward. Remember that this spot is the only one that does not move and that the movement of all other points is in a direct line away from this spot. Do not try to base your judgments on any other feature of the landing situation except the center of enlargement.
	Here is the next practice trial. Start at number 6 on your answer sheet and blacken the space underneath the letter corresponding to the spot at which you are aiming.
Title, "Trial 6," for 2 secs.	Trial Six.
Glide aimed at E, synchronized with voice.	Watch the pattern of radiation of the ground as you approach the runway. At the edges of the field this radiation is most noticeable, and if you follow its direction inward from the edges of the screen, all the directions converge at some point on the runway. This is the aiming point of your glide and here the ground seems to move directly toward you with no outward expansion. (Pause)
Screen goes blank for 6 secs.	Record your answer. Spot E was the aiming point of your glide. The space under E on your answer sheet should be blackened.
Title, "Trial 7," for 2 secs.	Trial Seven.
Glide aimed at spot D synchronized with voice.	Notice spots A and E. Are they both moving outward? (Pause) Are B and C moving outward? (Pause) Does spot D show any expansion? (Pause) Narrow your choice of the aiming point to the spot which shows no expansion. (Pause)
Screen goes blank for 6 secs.	Record your answer. (Pause) Spot D was the aiming point of your glide.
Title, "Trial 8," for 2 secs.	Trial Eight.

ACTION:	VOICE:
Glide aimed at spot B.	You are on your own from now on. You have 3 more practice trials.
Screen goes blank for 6 secs.	Record your answer. (Pause 3 secs.) Spot B was the aiming point of your glide (etc.).

Effect of Instructions in the Use of the Expansion Cue. In order to determine the effect of training by verbal instructions on test scores, the preliminary form of the test was administered to two groups of preflight students. The first group took the fifteen trials of the test; then was given a training period during which the items were shown again, each accompanied by specific instruction concerning the expansion cue and the correct aiming point; and finally was tested again with the 15 trials, in different order. A second group simply took a test composed of three sets of the 15 basic items. Off-centering of the correct spot was done in both cases. The means and standard deviations of scores of both groups on the test are shown in the following table:

Trials	*Trained group (N = 31)*		*Untrained group (N = 66)*	
	M	*S.D.*	*M*	*S.D.*
1-15	5.49	2.52	5.82	2.30
16-30	(Training)		5.82	2.48
31-45	7.17	2.37	5.89	2.14

These results indicate a genuine increase in score produced by the training with verbal instructions. In comparison, scarcely any increase of score results from giving the 15 items again to the group which was given no verbal instruction. The critical ratio between means of pre- and post-testing in the trained group is 4.46; between the means of the trained and untrained groups on the final 15 trials it is 3.49. The correlation between scores of pre- and post-testing in the trained group was found to be .11, while the correlation between the first and third administration in the untrained group was .37. For test-construction purposes, the results show the importance of verbal instructions to be (1) that they introduce ability to *learn the discriminations* into the score, and (2) that they bring about a desirable increase in average score on the test.

Construction of 16-mm. Form

Preliminary testing was carried out with groups of preflight pilot students in order to determine a proper distribution of the levels of difficulty of the test items to be incorporated in the test. Ten seconds of a given scene was fixed upon as being the best length for a test item. If this ten seconds was taken from the end of a scene, the selection of the aiming point proved to be easy, whereas if it was taken from an earlier point in the glide, the judgment proved to be difficult. It was expected that the speed of

flight represented in the scene would also affect the difficulty of an item taken from the scene, but this variable proved to be much less influential than the former. It was decided to utilize items taken from scenes at all three speeds.

The 15 basic scenes referred to above (after being offset upwards or downwards at random) were converted into 60 items. Two ten-second sections of film were taken from each basic scene. The last (easiest) 10 seconds of each scene are utilized in the first 30 items of the test. This was done by using each of the 15 basic scenes (cut to 10 seconds) twice. The difficulties of these 30 items range from 55 percent to 70 percent correct judgments. An earlier (harder) ten-second section of each scene was taken to make the remaining 30 items of the test, so chosen as to yield difficulties ranging from 45 percent to 55 percent correct judgments. The test thus consists of 60 items, 30 being easier, 30 harder, and each 30 being made up by taking the 15 original scenes twice. The first and last 30 items were arranged in a random order with respect to the 5 aiming points, the 3 speeds, and the two types of offsetting of the aiming point. The completed film consists of 10 practice trials, 30 items at a relatively easy level and 30 at a more difficult level. Judging from the samples of preflight students employed in the preliminary experiments, the mean level of item difficulty in the final test for similar samples should be not far from 55 percent. On the basis of incomplete and indirect evidence from the samples referred to, it is believed that test-retest reliability of this test will be satisfactory. Reliability coefficients obtained from a limited number of cases and distribution constants may be found in chapter 5. This test in 16-mm. form was administered to 1,200 students at Keesler Field. However, the results are not at present available for inclusion in this volume.

The test exists in the form of 16-mm. sound-film prints, each copy being accompanied by a mimeographed pamphlet on the nature and use of the test. It is 31 minutes in duration.

Potential Uses of the Test

This test of landing judgment has a number of potential uses. (1) It might prove to be useful for the selection of pilots as a measure of one aspect of landing aptitude. (2) It might be given as a type of proficiency test to measure a perceptual aspect of landing skill at different stages of training. (3) It could be used for training purposes. That a test developed essentially for classification purposes may have additional practical usefulness is not an ordinary occurrence. The feature of this test which makes it otherwise applicable is the fact that the instructions are, in effect, a training film on how to judge one's point of aim during an approach glide.

The appearance of the ground during an approach is not ordinarily familiar to a student pilot. A test of his ability *at the outset of training* to judge this appearance and to estimate where the glide would bring him down might not, therefore, be a fair test of his aptitude for making the estimate. In order to test his aptitude, the beginner must be given an opportunity to learn how to judge this appearance and how to estimate the point of aim. In an airplane, unlike all land-operated modes of locomotion such as automobiles and trains, the aim of one's path is not the axis of the conveyance. In the landing situation this discrepancy is of critical importance because of the limited degree to which a glide can be modified once it is established. The point of aim has to be determined by watching the ground, not the plane. Since this requirement tends to conflict with old locomotor habits, it is not easy to learn. The present test is set up to measure not original ability as such, but the aptness of the student in learning to make the judgments described after instruction and practice.

It has already been shown that the instructional part of the test produces an improvement in score. It might also be expected that actual training in landing a real plane would result in higher test scores, since the student becomes increasingly familiar with the appearance of the ground during landings. If the test can detect improvement in accuracy of landing, which remains to be determined, it may be used as a proficiency measure for a kind of *perceptual* achievement or skill supplementary to the performance scales already developed.

To determine the value of the film as a proficiency test, it can be administered to a group of primary school students on two occasions, once at the beginning of their course, and once again at the end of the course. An improvement in final test scores would be the result of either practice in the test or practice in making landings. If an improvement in score is found it will therefore be necessary to test a control group of students once in their last week of training. Scores for the control group which are comparable to the final test scores of the group tested twice will mean that the landing judgment test has detected a gain in actual landing proficiency. By correlating some flight criteria such as instructors' ratings or check ride ratings with test scores, the value of the Landing Judgment Test for proficiency test purposes could be obtained. This experiment could unfortunately not be carried out by the Psychological Test Film Unit.

Since the first ten-minute period of the test is devoted to instructions and practice in finding the motionless aiming point of the approach guide, the test may be used as a training film. The fact is that there exists no uniform method for teaching students what to observe in landing a plane. Analysis of the perceptual

situation during landings is not ordinarily made for students by instructors in a systematic way. Few, if any, instructors are skilled in the conscious analysis of perceptual cues, and few can point out the existence of the expansion pattern in determining the aiming point of the approach glide. There is, if this state of affairs exists, a need for an effective instructional film on the perceptual basis of flight. As one contribution to such a training film the present form of the Landing Judgment Test could be tried out. The correct answers for the different trials would have to be read off by the instructor immediately after the trials themselves. The test might, in short, be used for *practice* instead of for *testing.*

SUMMARY

If the human being had a single eye in the middle of his forehead, like Cyclops, the world of space would be registered by him in somewhat the same way as it is registered by the camera. The cues or retinal variables which constitute this registering of space on a two-dimensional retina, or film, have been elaborated in a theory which is more systematic than previous descriptions of "monocular" space perception have been. Individuals who look at a photograph perceive space with a surprising degree of accuracy as the photographic experiment has shown. They do so, moreover, without special training—without any more experience in the use of the monocular cues than an ordinary binocular individual possesses. These static stimuli for the perception of space are powerfully supplemented by the facts of retinal motion perspective, a theory of which was also elaborated. The motion picture screen can reproduce space, according to this analysis, with even better accuracy. The conclusion must be that the monocular basis of space perception is much more significant than has usually been recognized in psychology and aviation medicine. Research and tests in the field of space perception are needed with this conclusion as a starting point. Two tests of this type were described.

Evidence has also been presented to show that the objectivity of space perception at large distances is greater than the psychological experiments on spatial constancy would lead one to expect. Accurate perception at such distances is presumably necessary for high-speed flying. The actual performances of pilots in reacting to spatial cues far outstrip our understanding of the perceptual mechanisms by which they perform. An understanding is nevertheless required if the selection and training of future fliers is to be intelligently carried out.

CHAPTER TEN

The Instructional Techniques Peculiar to Motion Pictures*

INTRODUCTION

Although the motion picture medium was coming into wide use in the schools and colleges of this country before the war, it had its greatest application during the period of war training. The military services faced the problem of teaching a great variety of subjects and skills, some of them novel, technical and difficult, with few specialists, many students, and little time. In the effort to meet this situation a large number of training films were produced to supplement ordinary methods of teaching in the schools of the various military services. Some were made commercially, but the great majority were produced in military studios established for the purpose. The AAF Motion Picture Unit, staffed by professional writers, producers, directors, actors, cameramen, editors, artists, sound technicians, and laboratory technicians who were enlisted or commissioned in the army, was one of the largest of these studios. It produced training films for the specialized needs of the AAF. The output of this and other units was such that by the middle of 1944 the quantity of 16-mm. sound films approved and available for instruction in AAF schools was numbered in the hundreds. This figure does not include film strips for still projection in classrooms which were also produced in great quantity. The catalogue of approved training films and film strips published by the AAF Training Aids Division was a volume of 150 pages. The subject matter varied from orientation and morale films like "The Rear Gunner," dramatizing the importance of the aerial gunner in the combat crew, to technical films like "How to Rivet Aluminum." They differed widely in their method of approach to the subject and in their teaching value, and they were used in training with varying degrees of success.

As produced, these training films were capable of relieving instructors from a part of the burden of teaching, especially in technical subjects. The demand for them increased to a point

*This chapter was written by the editor. The research which is reported was carried out in large part by L. H. Borin (script evaluation) and C. H. Orvis (learning experiment). The conclusions stated were in most cases reached jointly with R. M. Gagne.

where it appeared that all courses on all kinds of subject matter would require training films. But whether these courses would profit by instruction utilizing the motion picture medium was unknown. Whether in fact the training films already produced were in all cases superior to ordinary methods of instruction was also unknown. On the basis of the slogan that "a picture is worth a thousand words," some producers of films were willing to go ahead and make pictures on any subject that could be taught. A difference of opinion arose on this question between those who saw no limits to the effectiveness of the motion picture medium and those who held ordinary methods of teaching to be equal or superior to it in the case of some subjects. The question in need of evidence was this: *For what subjects is the motion picture particularly adapted to give better instruction than ordinary methods provide, and for what subjects, if any, is it not so adapted?*

The problem was complicated by another difference of opinion over the type of motion picture which had the greatest instructional value. Professional film producers tended to put all possible ideas and lessons into dramatic form and to incorporate a story with characterization, humor, and dramatic incident wherever possible into a subject matter which, so it appeared to them, would otherwise be dull and technical. The entertainment value of an instructional film, it was argued, was something to be sought for in its own right. The critics of these films maintained that a straightforward expository or "documentary" approach would in many cases be superior and that the dramatic treatment of the subject sometimes was distracting and tended to run away with the picture. The argument in these terms could probably never be settled. In all likelihood there could be good or bad instances of dramatization and good or bad instances of exposition. The underlying question was this: *What are the instructional techniques peculiar to motion pictures which give them an advantage over other methods of learning?*

A great many experimental studies had been made before the war on the educational value of motion pictures in the schools. None of them, however, was relevant to either of the two questions just formulated. On the whole, they had been performed to demonstrate the effectiveness of the motion picture as an instrument of instruction, and had succeeded in doing so for the particular subject matter chosen and under the particular circumstances investigated. In the majority of these experiments the procedure was simply to compare the achievement of classes which did not include the feature of motion picture instruction with classes which did. Science, geography, history, and nature study had been popular subjects for experimentation, and considerable gains in achievement had been demonstrated. When examinations were

repeated after a lapse of time the superiority of pupils taught with films was usually even more marked than before. Little analysis had been made, however, of the specific ideas, insights, concepts, principles, and rules which had been better learned by the film-taught students than by the other students. And likewise there had been little systematic study of the differences, point for point, between the motion picture presentation of a subject and the equivalent classroom presentation of the same subject with the aim of discovering wherein the superiority of the film lay.

In order to make some approach to an answer to the questions formulated above, the Psychological Test Film Unit made an experimental study early in 1944 of the reasons for the superiority of a recently-completed training film. The film in question was agreed by all parties in the current discussion to be an excellent example of good motion picture instruction. It was concerned with a subject particularly appropriate for motion picture presentation and it employed techniques of which only the motion picture medium was capable. The study was undertaken at the suggestion of both the AAF Motion Picture Unit in Culver City, California, and the AAF Training Aids Division in New York City. The plan of the study was subsequently extended to include a complementary experiment on the reasons for the non-superiority (if such were the result) of another training film on a presumably *inappropriate* subject for motion picture presentation. This experiment was planned but for administrative reasons was never carried out. Some of the conclusions which are reached in this chapter regarding the instructional techniques peculiar to motion pictures are based therefore on incomplete experimental evidence. They rest on the implications of a single experiment. This evidence was supplemented, however, by the experience gained from analyzing the teaching methods of a large number of "shooting scripts" of training films awaiting production at the AAF Motion Picture Unit. Sixteen of these scripts were ones submitted to the Psychological Test Film Unit for evaluation of educational techniques in advance of production.

ANALYTICAL COMPARISON OF THE EFFECTIVENESS OF ALTERNATIVE METHODS OF INSTRUCTION

The training film selected for analysis had to do with the theory and practice of a newly devised system of aerial gunnery, called "position firing." The problem of the gunner in aiming at an attacking fighter from a bomber which was itself in rapid motion was difficult. It was necessary to lead the target, for example, not as one would lead an ordinary moving target, but by a deflection backward from the direction of the bomber's flight. The problem had been solved, however, in a systematic manner and the rationale

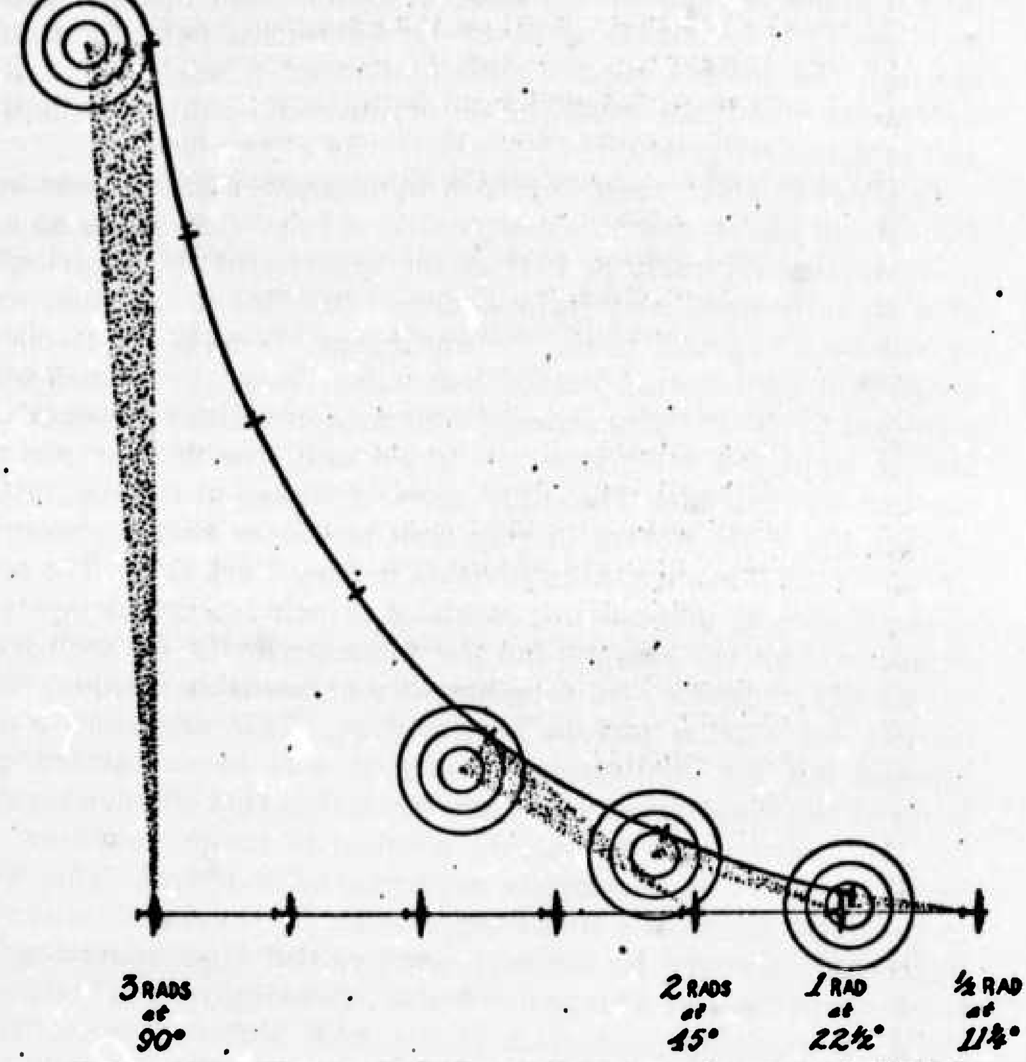


FIGURE 10.1.—Changes in Deflection During an Attack by a Fighter.

of the procedure, although difficult, was logical and could be thoroughly mastered by anyone having the necessary insights. The rules and procedures of firing were derived from the system and could obviously be learned more readily by understanding the system than by memorizing them in rote fashion. This state of affairs did not always apply to military subjects of study.

The practical importance of teaching this system to gunners as rapidly as possible was so great that all available methods of instruction were resorted to. Gunnery instructors were promptly indoctrinated in it, illustrated training literature was devised and distributed, and a training film was produced by the AAF Motion Picture Unit (TF 1—3366, "Position Firing"). The best efforts of the instructors, the writers of training literature, and the film producers had gone into this task. If the film proved to be a superior method of instruction, it would not be because of any lack of effort devoted to other methods.

Purpose and Method of the Experiment

The purpose of the study was in the first place to compare the overall efficiency of the training film with that of two other methods of teaching: an illustrated lecture (oral instruction with visual aids) and an illustrated manual (written instruction with visual aids). It was intended to analyze the subject matter taught, i. e., the system of position firing, into its basic rules, concepts and procedures, to make certain that all three methods of instruction, the film, the lecture, and the manual, covered these basic points or ideas, and then to take each point one by one and find out how well it had been learned by each method. Finally, it was planned to examine those basic points for which the film had shown its superiority as compared with those for which it did not, in the effort to analyze the instructional techniques which accounted for the success of the film.

The relative efficiency of a film, a lecture, and a textbook cannot be fairly compared unless each teaching method uses the best techniques possible for that method, and unless each covers the same basic points of the subject matter. The attempt was made in this experiment to match the three teaching methods for content and to make each method as effective as it was capable of being. A mediocre training film may prove to be more effective than a talk or a textbook if the latter are not supplemented by pictures, diagrams, and examples. But in this case what has been demonstrated is not the advantage of motion pictures, but simply of pictures. It was not intended in this experiment to measure the effectiveness of visual aids in general, but rather the unique effectiveness of the motion picture method of presentation. The lecture and the manual made use to an unusual degree of still pictures and diagrams; any advantage of the film would therefore be due to its unique characteristics, of which two are *motion* and *real action*. The three presentations were as follows:

a. Training Film 1–3366, "Position Firing," a 15-minute film produced entirely by animated photography, in which the subject matter was presented with commentary and animated diagrams in a logical order. The material was, however, also organized around a thread of story providing characterization and humor ("Trigger Joe").

b. A 50-page, pocket-size, loose-leaf illustrated manual entitled "Get That Fighter," employing advanced visual methods with a minimum of text and a maximum of diagrams, covering the same subject matter as the film. The diagrams were in color and the execution was extremely skillful. This manual was published by AAF Training Aids Division as of 1 November 1943. It was studied by the trainees without discussion or explanation.

c. A half-hour lecture on the same subject matter, organized

around a series of 19 lantern slides made from the booklet. The talk was written out for delivery in an informal spoken style, the diagrams on the screen being explained with a pointer. It was then revised, rewritten, and finally memorized for actual delivery. The talk included questioning of the class by the instructor to bring out salient points. The presentation was not pedantic, and used the methods of good face-to-face teaching. As revised and approved, it was judged by two experienced college teachers to be an example of excellent instruction. An abridged text of this lecture is given in Appendix A at the end of this chapter.

In order to compare the learning produced by these three methods, the subject matter to be learned had to be defined as exactly as possible, and an examination had to be constructed on this subject matter to measure the amount learned. The system of position firing was in early 1944 a fairly self-contained and systematic set of principles, concepts, rules and procedures for air-to-air combat firing by flexible gunners. Although not at all easily explained, the system was capable of being learned completely; once learned there was not much more to know about it except to put it into practice and to acquire skill. After study of all available sources, it was analyzed into 14 basic points, five of which were essential to an understanding of the system, and nine of which had to do with the operation of the system of practice.

It was then determined that the training film and the booklet covered these 14 points, and that no additional points were made by either which were believed fundamental to the system of position firing. The lecture was written especially to cover these points.

On the basis of the concepts and procedures embodied in these 14 points, as many test questions were written as it was possible to devise. A series of questions was adapted from an examination on position firing written by the Psychological Research Unit (Gunnery) then at the Central Instructors' School, Ft. Meyers, Fla. Other questions were originated as a result of the analysis referred to. Some of the questions were pictorial rather than strictly verbal. All items were of the objective five-choice type. The list of questions was then edited, revised and rephrased so as to eliminate poor misleads, verify the correct responses, and discard overlapping items. Twenty-five questions were finally selected for the examination. Each "point" of the system of position firing was represented by from one to six questions depending on how many different aspects the point had. For purposes of this experiment, it is believed that the examination represents 25 distinct items of knowledge about the theory and practice of position firing. The number of questions is relatively small but was the maximum number that could be devised without overlap, considering the limited scope of the system to be learned.

The corrected odd-even reliability of this examination proved to be .63 when administered. The mean level of difficulty of the items was in the neighborhood of 70 percent correct answers for the groups tested immediately after training, and 55 percent for the same groups tested two months later. The characteristics of this examination as a measure of ability to comprehend motion picture instruction are given in chapter 5.

The Experiment

The plan of the experiment was to teach position firing to three equivalent groups of beginning aviation cadets by the three different methods, and then to give the same examination to all three groups. In order for the mean scores of these groups to be measures of the amount of learning produced by the training, they had to be corrected by subtracting in each case a number representing the questions of the test which were answerable without training, i. e., by common sense knowledge and general information. A fourth equivalent group of aviation cadets was therefore given the test without any training whatever. The mean score of this group served as a base line for the other three scores.

In all groups, aviation cadets who had been given previous army training in gunnery or who had previous information about the system of position firing were eliminated from the experiment. All groups were given sufficient time to finish the examination. All groups were told to guess the right answer among the five choices for each question if they had any idea of it, and all scores were corrected for such guessing by scoring the test as number right minus one-fourth the number wrong. The trained groups were informed in advance that they would be given an examination immediately after the instruction.

In order to measure the effectiveness of the three methods for remembering as well as for learning, all four groups were recalled after a period of two months and were again given the examination. The cadets had no foreknowledge of this second examination and no access to information during the interval, so it may be assumed that forgetting occurred during this period in a normal fashion.

The instruction and the original testing were given within a one-hour class period to sections or classes of 40 aviation cadets. All were in classification squadrons awaiting entry into Preflight School at Santa Ana Army Air Base. Four hundred and eighty men participated, which reduced to 456 after eliminating those with previous information. Each of the four groups (*Film, Manual, Lecture, and No Training*) consisted of from 100 to 130 men selected at random from the total cadet population.

One superiority of the film over the other methods of instruction became evident in setting up the experiment. The film required only 15 minutes of running time, whereas it proved to be impossible to give an adequate lecture in that period of time, and also it proved to be too short an interval for adequate study of the manual. Since it had been decided to match the three methods in content and to compare the film with other methods *at their best level of effectiveness*, matching with respect to time was sacrificed. The lecture and the supervised study of the manual were each allowed 30 minutes instead of 15 minutes. The time allowed for the examination was 30 minutes in all cases.

Each section of cadets met in an ordinary classroom with a non-commissioned officer acting as instructor. This same instructor conducted all phases of the teaching. Trainees were told that they would have a lesson in a relatively new and important system of flexible gunnery. They were to learn as much as they could in the allotted time since an examination on the subject would be given them immediately afterwards. Nothing was said at any time about a subsequent repetition of the examination. A given section of cadets would then either be shown the film, or given the lecture, or the manuals would be passed out for supervised study. At the end of 15 minutes, or half an hour, the test was administered. The "no training" group was told in part the purpose of the experiment and that they were to "beat the examination" if they could. They could not expect to make good scores but were to make intelligent guesses whenever possible.

Results

Immediately After Instruction. Table 10.1 gives the mean scores obtained by the four groups on the examination. The fourth column shows the amount of learning produced by the three methods of instruction, i. e., the number of items of knowledge actually learned as a result of the training. This number is obtained by subtracting from the score of each group the mean number of items answered correctly by the "no training" group.

TABLE 10.1.—*Mean scores and amount learned immediately after training*

Group	*N*	*Mean Score*	*S.D.*	*Amount learned (score minus 5.42)*
Film (15 min.)	132	17.91	3.0	12.49
Manual (30 minutes)	101	15.43	4.2	10.01
Lecture (30 minutes)	101	15.19	4.6	9.77
No training	122	5.42	3.1	

It can be seen that the film produced considerably better learning of the material taught despite the fact that it occupied only half of the time required by the other methods. On the average, two and a half more items were comprehended—items which had been carefully taught by all three methods. The difference be-

tween the film and the other two methods is statistically significant (critical ratios of 5.13 and 5.04) but the difference between the manual and the lecture is not significant.

After Two Months. The film method yielded faster, easier, and better learning than did the classroom or textbook type of instruction, but how well was the subject matter remembered? Table 10.2 gives the mean scores and the amount remembered by the trainees after two months. It should be noted that the number of cases in each group who could be called back for retesting is reduced, especially in the case of the "no training" group. This reduction is due to transfers of cadets from the air base; there is no reason to suppose that the groups are not still representative samples.

The superiority of the film over the other types of instruction is just as marked for remembering as it was for learning, if one

TABLE 10.2.—*Mean scores and amount remembered after two months*

Group	*N*	*Mean score*	*S.D.*	*Amount remembered (score minus 6.96)*
Film	98	16.36	4.10	9.40
Manual	68	12.96	5.17	6.00
Lecture	86	13.58	4.93	6.62
No training	24	6.96	4.20	

compares the absolute size of the differences between the groups. The differences between the mean score for the film and the scores for the other two methods are again statistically significant (critical ratios of 4.5 and 4.1) but the difference between the manual and the lecture is not significant.

The mean number of items answered correctly by the "no training" group, who had nothing to remember except the examination itself, has increased from 5.42 to 6.96. Although the group is unfortunately small, this increase is probably real, since the second time one takes a test a better score may be expected merely by virtue of that fact. The base line for the amount remembered is taken therefore at 6.96 items. If the original base line were taken, the amounts remembered would be greater and the amounts forgotten would be less, but the relations between them would remain the same.

Good Students v. Poor Students. The training film evidently presented the subject matter in such a way as to promote faster and better learning by the average aviation cadet. But different individuals learn the same thing in different ways, some much more rapidly and efficiently than others. It is possible that the superior learners and the inferior learners among aviation cadets do not profit to the same extent when trained by the motion picture method, at least with respect to a specified achievement such as an examination in position firing. The best 30 percent and the worst 30 percent of each group were isolated on the basis of their exam-

ination scores and the same comparisons were made between these sub-groups that have already been made between the total groups. Of these sub-groups, all on whom data were available after two months were likewise compared, although the number of cases had further decreased—so much in the "no training" group that the "amounts remembered" become no more than rough estimates. The results are given in tables 10.3 and 10.4.

TABLE 10.3.—*Comparison of mean scores immediately after training of lowest and highest 30 percent of each group*

	Superior students				Inferior students			
Group	N	Mean	S.D.	Amount learned	N	Mean	S.D.	Amount learned
Film	40	21.28	1.27	11.28	40	14.28	1.57	12.71
Manual	30	20.13	1.17	10.13	30	10.30	2.38	8.73
Lecture	30	19.83	1.37	9.83	30	9.77	2.96	8.20
No training	37	10.00	1.75	....	37	1.57	1.96	

TABLE 10.4.—*Comparison of mean scores two months after training of lowest and highest 30 percent of each group*

	Superior students				Inferior students			
Group	N	Mean	S.D.	Amount remembered	N	Mean	S.D.	Amount remembered
Film	34	20.79	1.57	8.12	30	11.57	2.05	8.68
Manual	19	19.05	1.23	6.38	25	7.12	2.90	4.23
Lecture	24	19.79	1.53	7.12	24	7.75	2.15	4.86
No training	6	12.67	1.92	....	9	2.89	1.37	....

In these two tables the statistical significance of the differences between the scores obtained under the three methods of instruction may be judged from the critical ratios of these differences. They are presented in table 10.5.

TABLE 10.5.—*Critical ratios of differences between mean scores*

	After training		After two months	
	Superior	Inferior	Superior	Inferior
Between film and manual	3.91	7.94	4.47	6.45
Between film and lecture	4.38	7.58	2.43	6.62
Between manual and lecture	.91	.76	1.76	.86

Comparisons between amounts learned at different score levels are admittedly dubious unless there exists a scale of amount learned having equal units. This is not the case in the present experiment, since the items were not all of equal difficulty. It may well be that, had the examination been much more difficult, the results would have been different, but it is not easy to see how the examination could have been much more difficult since it covered the whole subject of position firing—a delimited and specific set of rules and principles capable of being learned completely.

Bearing in mind these limitations of the examination and of the subject matter, it appears that, while the good students did somewhat better when trained by the film than by the other methods, the poor students did much better. After two months the differences are as great as they were in the beginning. The high-scoring students have retained enough of what they learned so

that they still average about 20 right out of 25 items; among them the margin of superiority of the film over the other methods is slight. The low-scoring students achieved less on the examination and, among these students, those trained by the film remember considerably more than those trained by the other methods. The benefits of motion picture instruction in position firing, therefore, while significant for the good students, have their greatest application in the case of the poor students. The implications of this fact will be discussed later.

The tables suggest, although the differences are not statistically significant, that both the good students and the poor students *remember* the lecture slightly better than the manual, and therefore that, if any difference exists, face-to-face instruction is not lost as rapidly as undiscussed reading of a manual.

Reasons for the Superiority of the Film Method

In the attempt to understand why the training film "put over" the subject matter better and why it was remembered better, each of the 14 basic points of the position firing system covered in the examination was studied separately. The test questions under each point were listed and the percentage of trainees in each group who answered the question correctly was tabulated. Equality or superiority of the film to the other methods was evident in nearly all of the 14 points, with one exception. Out of the 25 items of information in the examination, 17 were known better by the film group than by the other groups. The exception referred to had to do with the rule for the direction which deflection ("lead") should take under certain conditions. These conditions were explained in more detail and more clearly in the lecture and the manual than in the film. Since the point *could* have been explained effectively in the film, but was simply neglected, the fact is emphasized that any method of instruction is bad if the content of the instruction is not adequate. The lecture and manual were checked to see if their content were equivalent to that of the film on all points where the film showed superiority. They were judged equivalent or in some cases better. The treatment was in some instances more detailed since the time (30 minutes) was greater.

Unique Advantages of Motion Picture Presentation. The concepts or points for which the film showed its *greatest* superiority were studied to see if they yielded any clues to its success. This analysis proved to be revealing. In comparison to concepts for which the film showed less superiority or none at all, the concepts which were most successfully taught were those which might be called "dynamic" in the sense that they deal with changing events or with the variation of one thing in relation to another. Examples are the concepts of increasing (or decreasing) amount of

deflection with continuously changing angle of attack of the fighter plane, and the requirement that the amount of deflection must be slowly changed while firing (20% to 40% superiority). Other ideas which were outstandingly better comprehended by the film-trained group seemed to be ones which get their meaning from *use* or *human action* and which are hard to describe in words or static diagrams. A "rad," for example, (the radius of the sight and the unit of deflection) was more clearly comprehended by the film-trained group who had seen how it was used and had *vicariously experienced* the action of sighting a .50 calibre machine gun, instead of having seen it diagrammed and described and having to imagine for themselves its actual use.

It is only reasonable that "dynamic" ideas should be more comprehensible when represented by film. They can be described only imperfectly in words, however skillful the speaker, and they cannot be represented except by graphs or curves, or specified except by formulae and equations. Learning such ideas by these means is difficult. But the variation of one thing with another can be shown directly with the motion picture and the grasping of the idea becomes easy. Likewise ideas of how to act or how to use something are frequently hard to describe in words or to represent in still pictures, since they are continuous in time, but they can be given explicitly with motion pictures.

The training film in question not only could but did show these "dynamic" concepts, such as the variation of deflection with changing angle of attack, and in fact did make the trainees experience the activity of using various mechanisms such as the sight. It is concluded that the superiority of this training film to the best alternative methods of instruction is due in large part to its success in teaching concepts and procedures which *motion pictures are peculiarly adapted to teach.*

In representing these dynamic concepts and courses of action, "Position Firing" makes use of a camera technique which is relatively infrequent in other instructional films. For long passages, the camera takes the position (literally the "point of view") of the trainee in the learning situation, seeing what he would see, rather than the more conventional position of an onlooker watching someone else in the learning situation. Nearly one-third of the time spent on instruction in the film is devoted to these passages. By this technique the student is enabled to experience the activity to be learned rather than merely the external features of the activity. He sees what it is like to do something rather than what someone else looks like when doing it. This expedient is possible, of course, only with motion pictures, or in the real situations, or in artificial representations of them. It is employed here in a way which is both thoroughly convincing and natural. It seems

probable that this feature of the film in question helps to explain its instructional success.

If it is true that "dynamic" concepts and courses of action are relatively difficult to understand, and that a subject matter based upon them is not easy to learn by ordinary methods of instruction, this fact helps to explain why the poorer students profited more from the film presentation than did the better students in this experiment. Part of the explanation is probably that the better students could learn nearly all of the system of position firing by *any* method of instruction, and they therefore profited less by a superior method of instruction. But this does not exhaust the question. The comprehending of "dynamic" concepts from verbal descriptions and motionless diagrams requires a very active imagination—the ability to translate static symbols into the complete reality which moves and progresses with time. The better students have an advantage in making this translation. But the motion picture prvides a short-cut. It can reproduce the moving reality directly and explicitly and the advantage which the apt have over the relatively inept may be reduced on this account. Verbal and symbolic thinking are not required. The results suggest the hypothesis that the effect of good motion picture instruction on the learning of certain kinds of specified achievements is to reduce the variability of the achievement among the trainees and to put them more nearly on an even footing.

Entertainment Value. Apart from such characteristics as motion and live action which the film possessed, there was another difference between it and the other methods of instruction. It made considerable use of humor, building up a character ("Trigger Joe"), and incorporating a thread of story. In other words, the film had what is often called *entertainment value.* Some of those in charge of producing training films for the military services would assume that this was an important if not the chief reason for the superiority of the film method of instruction. In all likelihood it was a significant cause of the success of the film, but there was no evidence in this experiment to show whether it was or was not. A division of the 14 basic concepts into those which were and were not pointed up by humor or dramatization was not possible to make. It was noteworthy, however, that in the case of "Position Firing" the humor and the dramatization were subordinated to the instruction and were used for the sake of instruction rather than for their own sake. In this film the ideas and rules of the subject-matter itself were the only sources of "Trigger Joe's" adventures and his final triumph; he was not allowed to run away with the picture. In some other training films, where the story and the humor served only to "dress up" the picture, it could be doubted that they had any instructional value whatever.

703318—47—18

Implications. Aviation cadets learned and remembered more about the system of position firing from a 15-minute training film than they did from a half-hour of either classroom teaching or study of a manual. Both of these latter methods made use of excellent visual aids, covered the same material point for point, and were believed to have been about as effective as it was possible for them to be. Analysis of the results suggested that the overall superiority of the motion picture in this experiment could be ascribed to the nature of the subject taught and to the use of instructional techniques possible only to the motion picture medium.

The subject matter taught was particularly adapted to motion picture instruction since it involved "dynamic" concepts and courses of action. The film explained these concepts and showed the courses of action directly without requiring complex verbal and symbolic thinking by the learners. A number of instructional techniques which were present in the film but which could not be utilized in the other methods of instruction may be listed:

a. The representing of how one thing varies with another in actuality rather than by graphs, diagrams, or verbal descriptions.

b. The showing of activities and the use of various devices as they occur continuously in time rather than as broken up into separate acts or parts.

c. The exploitation of the "point of view" of the camera for instructional purposes; the use of the camera in such a way as to enable the student to see what the learner would see in the real learning situation, and therefore to experience more nearly the activity to be learned. This is the "subjective" point of view of the camera which, although relatively little used in American films generally, is highly effective in the film in question.

It cannot be assumed that this experiment demonstrated the superiority of training films over other methods of instruction generally. The results prove only that the motion picture medium is potentially superior for certain types of subject matter and if certain types of instructional techniques are exploited. These subjects and techniques need to be more fully examined.

THE CHARACTERISTICS OF THE MOTION PICTURE AS A METHOD OF INSTRUCTION

The motion picture medium is not completely understood, nor have all its possibilities been explored, even by those whose profession it is to use it. The commercial film industry has grown so fast and has been so dominated by financial considerations that there has been little time for abstract thinking or research on the subject. The best directors and cameramen have been those who knew intuitively what could and could not be done, but they were

seldom able to state their knowledge explicitly. What they knew, moreover, applied to the use of the motion picture for entertainment and they were not prepared, when the Army's need arose, to adapt it in a thoroughgoing manner to the needs of classroom instruction.

Educators concerned with the use of 16 mm. films in the schools had made a beginning in the analysis of the motion picture as a method of instruction. But they were handicapped in making this analysis by the existence of a fundamentally unanswered question in the field of visual education: were sound films to be produced as a *method* of training in which the functions of the teacher are, for the time being, reduced to zero, or were they to be produced as an *aid* to instruction requiring the active participation of the teacher? If the first alternative were chosen, an educational film would be designed as a self-sufficient teaching session, covering a topic of considerable scope and having a beginning and an end of its own. If the second alternative were chosen, a film would be produced of limited scope and of little value if shown by itself which could be subordinated to the teaching of any instructor and fitted into his own teaching plan. In the first case, the film is a substitute for an equivalent amount of ordinary classroom teaching. In the second, it is an aid to teaching of the same type as the lantern slide or the blackboard, except more lifelike. The disadvantage of the first is that there is a loss of the personal relationship between students and teacher. The difficulty with the second is that the use of films in this way would be laborious for instructors, and it would require relatively novel procedures in the employment of a projector and the illumination of the classroom. The procedures have been discussed in Chapter 4, and the advantages of this use of motion pictures in teaching were pointed out. The fact was that, except for motion picture tests, no training films of the latter type were produced for use in Army schools during the war. The films produced were self contained and self teaching. They were, therefore, in a practical sense, an *alternative method* of giving instruction—a substitute for an equivalent amount of classroom time.

It was in the light of this situation that the experimental comparison was made of alternative methods of instruction. If training films do in fact take the place of a certain amount of classroom teaching, despite the protests of visual educators that they were never intended to do so, then it becomes necessary to show that in a given instance the film will be more effective than the teaching it displaces. In some instances this will be so; in other instances it may well not be so. Whether it is or is not will depend on, first, the subject matter taught and whether it is amenable to motion picture presentation and, second, whether utilization has been made in the film of those techniques of instruction which ex-

ploit the possibilities of the medium. These are the practical problems which face the makers of educational films—the appropriate selection of subject matter, and the proper use of the opportunities which the motion picture affords for a superior type of instruction.

The Educational Techniques Available in Films

In the course of examining a considerable number of shooting scripts of AAF training films for the teaching methods which they employed, a tentative list of general techniques was gradually evolved by personnel of the Psychological Test Film Unit. Partly on the basis of the experiment already reported and partly from experience gained in constructing psychological tests on abilities which motion pictures were especially fitted to uncover, the list was narrowed down to a special set of techniques which seemed to be employable with motion pictures but not so readily possible, or even impossible, with ordinary methods of instruction. These methods or characteristics of motion pictures will be discussed.

The Overcoming of Spatial and Temporal Difficulties to Comprehension. The motion picture screen can give an enhanced representation of three dimensional space, because it can show motion-perspective (Chapter 9) and therefore can sometimes clarify the spatial relations of objects. By animation or special effects photography, it can juxtapose events never otherwise seen together. It can enlarge movements and objects too small to see by ordinary vision, and it can reduce the size of movements and objects too large to see under ordinary circumstances, such as storms and cloud formations. It can enable the learner to see through solid obstacles, and inside of operating mechanisms. The construction, by animated photography, of moving cross-sections and of completely instead of partially visible cutaway models was one of the accomplishments of wartime training films. By animation, it is possible to highlight and emphasize certain parts or processes of a situation and to direct the attention of the student in a way appropriate to the verbal description being given by the unseen narrator. In many ways the motion picture can "visualize the invisible." Of equal significance is the fact that by animation, or by slow-motion or time-lapse photography, it is possible to modify the velocity of a series of events so as to make them comprehensible. The motion of a dropped bomb, for example, can be followed slowly and discussed; it can even be transposed or reversed.

The Direct Representation of Events in Time. The study of the training film on Position Firing showed that "dynamic" ideas and concepts were better grasped by the film-trained group of students. The motion picture can represent how one thing varies with

another directly. It is superior in this respect to graphs, plots, verbal descriptions, or algebraic functions. The concept of simple harmonic motion, for example, can be grasped from an animated diagram in a moment, whereas a static diagram would require study and an active effort of imagination.

The sequence and pace of what the student is learning to do can also be shown directly on the screen. The performance he is striving for is usually something which occurs continuously in time, whether it be the firing of a .50 calibre machine gun or the use of the simplest tools and mechanisms. This tempo can be shown the student in advance of actual practice.

The Showing of Situations From a Subjective Point of View. In the learning of position firing, as in other types of subject matter where the visual aspects of the achievement are prominent, the evidence suggested that when the motion picture camera took the position of a *performer* instead of the more usual position of an *onlooker* in the learning situation, the student could experience the task for which he was being trained as it appeared "in life." He cannot, of course, get any kinaesthetic sense of how it feels to respond, but he can to some degree see responses being carried out. The motion picture camera can, for example, show hands in action; it can represent motion of the learner or locomotion from place to place; it can draw back or move forward or shift the direction of view as a performer would shift his position and the direction of his eyes; it can take a priviledged point of view of the scene. The imitation of "being there and doing it" is of course imperfect, but a surprising degree of realism is possible for a skillful animation artist or cameraman. The student who looks at the screen tends to become a participant in the action represented.

Vicarious Practice or Reinforcement of Right and Wrong Acts. Most important of all instructional techniques, because it is so fundamental to the learning process, is the capacity of the screen to portray the consequences of correct and incorrect actions. The student absorbed in a motion picture can go through the experience of committing all the errors he is ever likely to make, and in the end experience a kind of vicarious triumph by seeing the problem solved or the task done in the right way. He can make all the wrong choices and then the right choices, in advance, and without risk, in each case his reaction being differentially reinforced by the visible failure of the wrong ones and the success of the right ones.

The operation of vicarious reinforcement in motion pictures depends on the tendency of the student to participate in the action. It can be enhanced by the subjective use of the camera as described above. But it can also be induced in quite a different way—one which is more familiar to moviegoers in general. A dramatic story

also tends to make the onlooker participate in the action if its protagonist is someone with whom he tends to identify himself. The adventures of this character are sympathetically experienced by the moviegoer or the student. His mistakes and his triumphs are vicariously felt—sometimes with great vividness. It is probably in this broad sense that fictional movies may be said to teach, whatever the "truth" of the teaching. Consistent use was made of this dramatic method in the training films made by the AAF Motion Picture Unit. The hero was portrayed as forgetting to use oxygen on a flight at high altitude, losing consciousness, and coming close to disaster; or failing to make a preflight check of his airplane and crashing; or falling in with a girl under romantic circumstances and through carelessness becoming diseased; or firing with improper deflection at an enemy fighter and getting no results, but later firing correctly and seeing his enemy go down. The dramatic build-up to these incidents was frequently elaborate, and the consequences of doing the right and the wrong thing were shown with vividness and realism. The element of adventure and dramatic incident was incorporated in so many of these pictures that they were criticized by those who held that exposition and non-fictional treatment were in some cases preferable. The argument, however, was usually carried out in terms of whether the *entertainment value* of such films was an important consideration. *It may be suggested that the real value of dramatization in educational films is not entertainment as such, but the opportunity it provides for identification with a protagonist and for participation in his failures and successes.*

Personalization of Abstract Ideas. The motion picture offers even greater scope for cartooning and caricaturing, especially in the case of animation photography, than do other media. Commercial films are full of stereotypes, the actors themselves become "typed," and everyone is familiar with the animated type-characters of Disney. Such cartoon characters find application to education films in being used to represent or stand for abstract ideas. The concept of "lift," a force which maintains an airplane in flight, or the idea of an electron as it flows through a circuit, may be personified by animation and the action of the abstract force or of the theoretical particle may be shown by humanized actions of the cartoon characters. The "lift" is seen pushing up on the wings and the "electrons" are seen scampering around the circuit. It can scarcely be questioned that the cartooning of these ideas makes them memorable, and that the humanizing of the action makes it understandable. But the cost is frequently oversimplification of the abstract ideas and the action. Action and force are also sometimes visualized in animated training films not by personifying them but by devices such as pulsing arrows,

symbolic bolts of lightning, and the like. Such techniques can be, at their best, vivid and impressive, but the learning which results may well be insufficiently abstract.

Comic Emphasis. Teachers are aware of the value of comic illustrations and humorous incidents in ordinary teaching; all that the film can add is dramatic portrayal of the situations and characters. Since commercial film makers have a great facility at such portrayal, it was to be expected that the wartime films would utilize comedy wherever possible. They did so, frequently to the extent of being entertaining or amusing to persons having no interest whatever in the subject matter of the film. Sometimes, as in "Position Firing," the comic misadventures were fitted to the needs of instruction and emphasized a point which ordinary exposition probably could not have done. In too many other films, however, the humor was introduced simply for "comic relief," and was therefore extraneous to the instruction. At its worst it was merely "cute." Comic emphasis is probably a genuine technique of instruction but comic relief is at best a principle of showmanship.

Weaknesses of the Motion Picture Medium for Instruction

There are limitation to the motion picture as a teaching method as well as opportunities. Assuming that a film is to be presented as a teaching session with the voice of the narrator temporarily taking over the function of the classroom teacher, the comparisons that can be made are not all in its favor. It is sometimes claimed that students are better motivated by and more interested in a film, but this claim is meaningless since the result will depend on how stimulating the teacher is with whom the film is compared. With respect to the opportunity for repetition, motion pictures have no greater capacity than classroom teaching—probably less—for promoting learning. An instructor can go over the same material in different ways, but the film cannot. With respect to the optimum distribution or spacing of learning periods the film is in the same category as classroom sessions. With respect to the organization of the material to be taught, the educational film writer and the classroom teacher face the same problem, and neither has any advantage over the other.

In certain respects, the film is probably definitely inferior to classroom teaching. It involves, as ordinarily used, a loss in the personal relationship between student and teacher—a loss of the active attitude which a good instructor can elicit from a class merely by standing in front of it. With respect to the opportunity for discussion, for asking questions, and for the personal give and take between student and instructor, the film is at a considerable disadvantage. With respect to activities by the students such as writing, taking notes, and making responses to questions, the film

is also at a disadvantage, although this could be mitigated to some extent if the films were designed to do so.

Implications

The motion picture medium possesses opportunities for teaching which are superior to ordinary methods or even not possible for them. It also possesses limitations and weaknesses as compared with them. If educational films are to be produced to compete with classroom teaching, as seems likely, the opportunities should be known and explored, and the limitations should be borne in mind.

AAF training films during the war were for the most part written in the form of scripts by professional screen writers and supervised by motion picture producers. The instructors who were to *use* the films were represented only by a "technical adviser" for any given production, the title being the same as was customary in commercial studios. The experience which the writers brought to this task was principally gained from the writing of screenplays. This situation was inevitable, since script writing was an art with which few educators were familiar. But it had the effect of making the training films somewhat more like screenplays and somewhat less like teaching sessions than they would otherwise have been. The differences of opinion over entertainment value, dramatization, or exposition, and the use of comedy arose from this situation.

The implication of this situation for future production of instructional films is that teachers and screen writers need to collaborate. The teacher needs to know something about the writing of a motion picture script and should learn to visualize what can and cannot be shown on the screen. The screen writer needs to know something more about the practical problems of teaching and to acquire a realistic sense of how much he can count on the spontaneous attention of a student audience. Both need to understand the opportunities and limitations of the medium as a teaching method.

Appendix A

CLASSROOM INSTRUCTION (LECTURE) ON POSITION FIRING

[Lantern slides accompanying text are indicated by the numbers from 1 to 19.]

You are here today to listen to an illustrated lecture on position firing. Pay close attention to what is said because at the end of the discussion you will be given an examination to determine how much you have learned. Furthermore, at some time during your cadet training, if you plan to be either a bombardier, navigator, or fighter pilot, you will have to learn this subject in greater detail, and you will benefit at that time by whatever you can remember of what is said this (morning—afternoon.)

This will be the story of "position firing"—as far as we know it now. The most interesting chapters are still to be written by flexible gunners in combat. The need for the system arose as questions in the minds of combat air crews regarding their own inability to shoot down enemy fighters and in the minds of those charged with the responsibility of training them. Here is how "position firing" answers those questions.

.. ..

What is "position firing"?

Position firing is a system by which an aerial gunner is able to take and maintain proper deflection (lead) in firing at an enemy fighter. In other words it is a method of aiming which takes into account such factors as the forward motion of the plane in which the gunner is situated, the motion of the target, and the flight path of attacking aircraft.

Is position firing a complete firing system?

No, it is not. Position firing assumes that the enemy fighter is following a curve of pursuit on the gunner's own ship. Practice in position firing will develop the ability of the gunner to deal with fighters making attacks on his own ship or element only. Fortunately, it is under such circumstances that the fighter is easiest to hit, because the pursuit curve is a predictable path. Also the fighter is not dangerous to the gunner's ship or element; i. e., cannot fire on it with much chance of hitting it consistently, unless he is following a curve of pursuit. It is the estimate of combat gunners that an overwhelming majority of enemy fighters knocked down are following such tactics. The other flight paths which the

fighter may follow while not attacking the gunner's ship are so varied that they are practically beyond the ability of the very best gunner to hit. So we needn't be concerned with them.

Every enemy fighter is dangerous and needs watching, but he becomes really dangerous only when he starts a direct attack on you. This is also, however, the time when he's easiest to hit. A fighter plane has fixed guns, all facing forward. To aim his guns, the pilot must aim his plane. The fighter coming in to attack you must keep aiming at the spot where you will be by the time his bullets get there. If he aimed directly at you, he would miss completely because by the time his bullets arrived at the spot where he's aiming, your plane would have moved on ahead and the bullets would pass behind you. The exception to this statement, as will be mentioned again later, is when the fighter comes at you either *directly* from behind or from *directly* ahead of you.

Let's look at an illustration. (1) In order to keep aiming at the spot where you will be by the time his bullets get there, (in other words, to maintain the correct lead over a period of time) the fighter must fly in a curved path, called the pursuit curve. As he flies along this curve, he appears to slide in toward your tail. (Point out on slide.) He *must* follow this pursuit curve, which is a *predictable* course, in order to hit your aircraft consistently. A fighter plane following a pursuit curve can be recognized by the fact that he will appear to you to be coming at you head-on while at the same time sliding toward your tail.

Now let's consider in detail the business of aiming at an attacking fighter plane. Believe it or not, when a fighter is making his attack, you don't aim ahead of him as in most other shots. You always aim a little bit *back, between him and the tail of your own plane,* because the forward speed and direction of your plane is imparted to the speed and direction of your bullets. Look at this diagram. (2)

The plane is moving to the right at a speed of approximately 225 m. p. h. If you aim here along this black line, your bullet will take this direction, indicated by the red line, immediately upon leaving the muzzle of your gun, owing to the forward motion of your plane. Remember, a bullet shot from a moving plane keeps the forward speed of that plane as well as its own velocity, and its direction, too, is in part the forward direction of your plane as well as the direction in which you aim. Another diagram (3) will help to make this entirely clear. The direction of the plane is forward and to the right. In every case, the black lines indicate the direction of aim; the red lines next to each indicate the actual course which the bullet follows. Study the four firing situations represented in this drawing; and make sure that you understand thoroughly the principle involved. (Pause.)

If you make the mistake of leading *ahead* of where the fighter is pointing while he is shooting at you, allowing for *his* forward motion on the pursuit curve, you will miss, because you did not allow for the forward speed *your* plane gave to your bullet. Here's what would happen. (4)

The fighter is coming in here on his pursuit curve. You are going along in this direction (to the right) in your bomber. If you lead ahead of him here in your aim, along the black line, your bullet will travel *way* ahead here along this red line, owing to your own plane's forward motion, and you will miss completely.

The way to allow for your forward speed is to aim *back* between the attacking fighter and the tail of your own plane at a point on the line along which the fighter slides toward your tail. The amount by which you aim behind is called *deflection,* and you aim back toward your own tail, relative to the position of the fighter, whether he attacks from above, below, at your side, or even toward the front.

Look at this drawing of an attack from the side. (5)

The fighter is coming in at you from the left and to the rear. It might seem that you should aim ahead of him so that your bullets will intercept him. But this is not true. Instead you aim behind him toward the direction of the tail of your plane--along the black line. And then, because of the forward motion of your plane, the bullet goes off here along this red line and intercepts the fighter at this point. The same would be true if he attacked from above or below.

Look at this diagram of a plane coming in from overhead. (6) You aim back a little in the direction of the tail of your own plane, in order to hit him here. (Point.) Remember that this principle is an unvarying rule, when shooting at a fighter coming at you in a pursuit curve. (Show 7 and explain.)

A word about trail and gravity. It is, of course, true that trail, which is the drag of the air on the bullet, and gravity do exert a deflecting influence which is important for some shots, but not nearly so important as the correction you must make for your own plane's forward speed. All the deflections that will be mentioned in a few minutes have been corrected for trail and gravity.

Now, let's see exactly how you must aim to get the fighter before he gets you. The amount of the deflection that you make in your aim depends upon the direction from which the fighter attacks. You must learn to recognize just 4 key directions which the fighter may assume and then learn the proper amount of deflection for each. (8) Each of these key directions is at a certain angle to the line of flight of the bomber, either 90°, (point) 45°, (point) 22½°, (point) or 11¼°. (point). You will notice that each angle is

exactly half the size of the next largest one. But these four directions are not just simple straight lines like forward, backward, above, or below. Because your aircraft is suspended in midair, any forward or backward direction may be considered to include a circle or cone that goes completely around your plane. Look at this illustration (9) and I think it will be clear to you. It is the same as the one you just saw, except that the four directions are continued all the way around the ship. These cones and the circle around the middle of the plane may be numbered, and used as your four key directions in aiming. The center one is numbered 3, these two (which are similar but opposite) are 2, these 1, and these ½. The numbers refer to the amount of deflection you must allow in aiming at fighters approaching from each particular cone of direction. You give all fighters attacking from the surface of the same cone the same deflection. (10) All three of these fighters (point) are on direction three relative to your bomber, therefore, you give each one a deflection of 3, back toward the direction of your own tail. Both of these fighters are approaching on cone 2; therefore you would give either one of them a deflection of 2 in aiming, back toward the tail of your own ship. Let's look concretely at some examples of just exactly which direction your deflection would take. Let's work on direction three—the three planes on the white circle of attack direction. How about this top one? Which direction should your deflection take? Would you aim above him, below him, or to one side of him in order to allow for the forward motion of your own plane? Anybody?........You should aim to the left of him on this drawing, because your rule states that the deflection in your aim must be back along the line which the fighter follows in sliding in toward the tail of your ship and in the direction of your tail. Also you can see that the forward motion of your plane is going to throw your bullet this way (show); so you must aim in the opposite direction, back this way (show) or to the left of him to allow for it.

Now, how about this one here on the side? Picture yourself here substituting for the waist gunner. The fighter is flying toward you. On which side do you aim? Anyone?........You aim to the *right* of him (remember, we're speaking from your point of view, watching from the bomber). In other words, you aim here (point), back toward the tail of your ship.

Take the one below for a last example. You aim here (point), to allow for the fact that your own forward motion is going to throw the bullet forward this way (show) to hit him. The same holds for this plane in cone 2. You aim above him—toward the direction of your own plane's tail. Or, for this one coming from behind, you aim below him—here. (Explain straight-ahead and astern shots at this point.)

Two more problems must be considered: first, how to recognize the four key directions, and second, the *amount* of deflection to be used in aiming at fighters approaching from each of those directions. (9) One of the directions—the one that we have designated as number 3—can be used as a standard by which to estimate the others. This direction is a circle surrounding the plane at a 90° angle—straight out, above, below, to the sides, or in *between* any of these, around this central circumference.

Direction number 2 is halfway between *number three* and the *nose* or *tail* of your plane, or an angle-off of 45°. Number one is in turn halfway between number *two* and the nose or tail of your plane, or an angle off of 22½°. Lastly, the direction designated as ½, this one here (point), or here (point), is halfway between number one and the ends of your plane, or an angle-off of 11¼°. These can be remembered easily because each is just one-half of the next larger one, with 90°, the largest one, as a base point from which to estimate. In actual practice, you have to get a sort of feeling for these positions or directions, since it is impracticable to remember them in terms of the number of degrees of the angle-off, that is, the angle of approach of an attacking fighter.

Now, finally, we must consider the *amount* of deflection used in aiming at fighters coming in from these key directions and how this amount is estimated.

(11) In your sight, the deflection is the horizontal distance between the bead or pipper in the center and the fighter. (Show) The amount of deflection is measured in rads (a contraction from the word "radius"), which is simply the distance from the pipper in the center to the edge of the ring. This drawing shows clearly what is meant by rad and deflection.

You must be able to estimate mentally the number of rads from the center-bead of your sight to the enemy plane. In this drawing the deflection is *three* rads. This is the amount of deflection you would use when the fighter is coming in from cone 3, the central circle about the plane. You must also be able to estimate two, one, and one-half rads, the amount of deflection to be used for planes approaching from cones 2, 1, and ½ respectively, either front or rear.

The following drawings show the exact deflection for each of the 4 cones. The success of the mission and the lives of the crew depend on whether you spot the cone from which he attacks and use the correct deflection. (Show 12, 13, 14, and 15 with appropriate explanatory comments.)

You must keep in mind that these cones are not actually separate, discreet intervals in space. The fighter does not, of course, jump from one cone to another, but rather moves continu-

ously from one to another. When a fighter flies an attack curve, he always slides in toward your tail. (16)

A rear attack begun on cone 2 (point) moves toward cone 1. Front attacks from cone 1 move toward cone 2, etc. When you make your deflection back toward the tail of your plane, it is along this *line of apparent motion* of the fighter.

As the fighter slides from one cone to another, you must *adjust* your deflection toward the value for the *new* cone.

When the fighter comes in range (about 600 yards), start firing with a 2-second burst. To allow for the fact that the fighter moves from one cone toward another during this time, you must let the fighter drift in your sight ½ rad toward the bead or pipper for attacks *behind* the beam, and ½ rad away from the pipper for attacks forward of the beam.

Here's a picture of what happens on an attack behind the beam. (17) Notice how the gunner allows the fighter to drift in *toward* the central pipper on his sight as the fighter moves from one cone toward another, sliding in toward the tail of the bomber.

After an initial 2-second burst of fire, check your aim and deflection and then fire another 2-second burst. By this time the fighter will probably be in his break-away, having finished his attack. Or better yet, by this time he will be going down in flames.

The number of rads deflection which you should use is not affected by range. If you need a three-rad deflection to hit a fighter at 600 yards, you need the same three-rad deflection at 300 yards, as long as the fighters at these two different ranges are on the same cone, or angle of approach.

Look at this drawing (18). You use a three-rad deflection for a fighter on this cone (which is a 90° angle of approach), whether he's out here at 600 yards (point) or in here at three hundred. (Point). This is based on the simple proposition of similar triangles. The angle is the same, whatever the range, as long as he stays in the same cone. Therefore, the deflection is the same. All these deflections that have been mentioned apply to a 225 m. p. h. true air speed for the bomber and 325 m. p. h. for the fighter.

Remember these two simple rules:

First, when the fighter has started his attack on you, aim a little backward between him and the tail of your own plane, to allow for *your* forward motion.

Second, spot what cone he is on, and use the correct amount of deflection for this angle of approach. (Slide number 19)

The success of the mission and the lives of the crew may at some time depend on your gunnery. If you learn and use the simple directions presented in this discussion of position firing, you can depend on your ability to shoot from the sky any fighter which attacks you.

Appendix B

INVENTORY OF PSYCHOLOGICAL TEST FILMS

[16-mm. sound prints on file at Department of Records and Analysis, School of Aviation Medicine, Randolph Field, Tex., and in the Office of the Air Surgeon, Hq. AAF, Washington, D.C.]

Name of test	Running time (minutes)	Film project No.	Location of negatives
Aptitude tests			
Estimation of Velocity Test CP205B-I	22	....	Motion Picture Branch, Air Technical Service Command, Wright Field, Dayton, Ohio.
Identification of Velocity Test CP205B-II	17	....	MPB, ATSC, Wright Field, Ohio.
Est. of Rel. Velocities Test CP205B-III	19	....	Do.
Landing Judgment Test CP505E	28	112A	AAF Motion Picture Unit, Culver City, Calif.
Distance Estimation Test CP212A	(Still photographs)		Dept. of Records & Analysis, School of Aviation Medicine, Randolph Field, Texas.
Flying Orientation Test CP107A	18	....	AAFMPU, Culver City, Calif.
Landing Orientation Test CP106A	(35 mm.)	112D	Do.
Minimal Movement Test CP213C	20	112E	Do.
Drift Direction Test CP221B	19	112F	Do.
Flexibility of Attention Test CP411E	15	....	MPB, ATSC, Wright Field, Ohio.
Integration of Attention Test CP415A	16	....	Do.
Successive Perception Test CP509C-I	14	112B	AAFMPU, Culver City, Calif.
Successive Perception Test II CP509C-II	24	112C	Do.
Plane Formation Test CP805C	15	....	MPB, ATSC, Wright Field, Ohio.
Motion Picture Comprehension Test CI625A	....	....	No negative in existence.
Proficiency Tests			
Aircraft Recognition Proficiency Test (Preflight Level)	24	....	No negative in existence.
Aircraft Recog. Proficiency Exam. (Form A) (TF 1-3368A)	24	1735A	AAFMPU, Culver City, Calif.
ARPE (Form B) (TF 1-3368B)	24	1735B	Do.
ARPE for Flex. Gunners (A) (TF 1-3392A)	23	1898A	Do.
ARPE for Flex. Gunners (B) (TF 1-3392B)	23	1898B	Do.
Navigation Proficiency Test (Map Reading & Dead Reckoning)	90	9125	Do.
Target Identification Test for Bombardiers (Preliminary Form)	25	209	Do.

☆U. S. GOVERNMENT PRINTING OFFICE: 1947—703313